

AC AAQ51543;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 24-MAY-1994 (first entry)  
 XX  
 DE Lipopolysaccharide induced protein gene.  
 XX  
 KM Macrophage; induced; lipo-polysaccharide; antitumour;  
 KM antiinflammatory; trypanocidal agent; antibody; cell proliferation;  
 KM activation; cytotoxicity; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1810..3315  
 FT /tag= a  
 FT 1980..2228  
 FT /tag= b  
 FT /note= "Intron 1 is only partially sequenced  
 (represented by (N)40 in the sequence  
 estimated length ca. 5400 bp"  
 FT intron 2615..2846  
 FT /tag= c  
 FT /note= "Intron 2 is only partially sequenced  
 (represented by (N)40 in the sequence  
 estimated length ca. 7900 bp"  
 FT intron 2907..2995  
 FT /tag= d  
 FT /note= "Intron 3 is only partially sequenced  
 (represented by (N)40 in the sequence  
 estimated length ca. 1000 bp"  
 FT polyA\_signal 3841..3846  
 FT /tag= e  
 XX  
 PN WO9322437-A1.  
 XX 11-NOV-1993.  
 XX  
 PD 28-APR-1993; 93WO-EP01022.  
 XX  
 PR 30-APR-1992; 92EP-0401231.  
 XX  
 PA (INNO-) INNOGENETICS NV SA.  
 XX  
 PI Devos K, Franssen L, Van De Voorde A, Van Heuvelwyn H,  
 X  
 WPI, 1993-368796/46.  
 X  
 P-PSDB; AAR43682.  
 XX  
 PT New polypeptide induced in macrophage(s) by lipo-polysaccharide -  
 PT useful e.g. as antitumour, antiinflammatory or trypanocidal  
 PT agent, also related nucleic acid, antibodies, anti-sense cpds.  
 PT etc.  
 XX  
 PS Claim 9; Fig 1; 108pp; English.  
 XX  
 CC The polypeptide induced in macrophages by lipopolysaccharide  
 CC stimulates cell proliferation (esp. when costimulated with IL-4)  
 CC promote activation, cytotoxicity, and mobilisation of LAK cells;  
 CC promote recruitment of suppressive peritoneal exudate cells;  
 CC promote generation of immunocompetent lymph node cells (LNC) and  
 CC have trypanocidal and trypanolytic activity. The human and murine  
 CC sequences are given in (AAQ51543-45), peptide fragments able to  
 CC generate antibodies are given in (AAR51951-61)  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 SQ Sequence 3861 BP; 542 A; 1241 C; 1292 G; 658 T; 128 other;  
 Query Match 15.1%; Score 69.2; DB 14; Length 3861;  
 Best Local Similarity 53.1%; Pred. No. 0.0015;  
 Matches 170; Conservative 0; Mismatches 148; Indels 2; Gaps 1;  
 35 CAGCCCTGCGCTGGCCAGAGCGCGAGCGCCGCTCGGTGAGACTGGGGGTG 94

Db 2086 CTGCCCCGCGCCCGCCCCCGCGCCCGCACGCTGCGCGCCCGGAGCGAGAGG 2027  
 QY 95 GAGGTGCCCCGAGAGCTACCCAGCGCCGGAGTACTCCCGCTCAACACTCTGGAGCTGCAGT 154  
 Db 2026 GACGCGAAGAGGAGGGGGCCGGGGTCCGCGCCGACACTACCTCCCTTCACAGC 1967  
 QY 155 TCCCTGGGTGGCCCGCGAGAGCTGAGCCCGGGCTGAGAGGATGGCGGGGCGGAGCGAGG 214  
 Db 1966 TGCACCGGTGCTGAGATGATGTCGGCGCCCGCGCCCGACAGAGCCCGGACAGAGAGAG 1907  
 QY 215 GCGGGGGCGGAGCTGTCACGTGAGAGAGCGCGCGGGGCGGAGCGGGGCGGGGCGCGCG 274  
 Db 1906 GCAGCAGCGGGAGCGCGCGCGCGGGGCGGGGCGGGGCGGGGCGGCGGCGGCGGCGGCG 1847  
 QY 275 CCGGCTCTTAAAGCCCGCGAGCGGCGGAGCGGCGGAGCTCTGTGCGCGGAGCGGAG 334  
 Db 1846 CCGCGCGCCCGCGAGCGCCCGCGCGCC--GCGCCCGCATGCTGTGCGCGCGCTCCGCG 1789  
 QY 335 GCCCGGGGAGACGCGGAGC 354  
 Db 1788 CCGGCGCCCGGCTCCGCGAGC 1769  
 RESULT 14  
 ID AAX53491 standard; DNA; 114955 BP.  
 XX AAX53491;  
 AC AAX53491;  
 XX  
 DT 05-JUL-1999 (first entry)  
 XX  
 DE Human adenosine A1 receptor antisense oligonucleotide fragment.  
 XX  
 KW Antisense oligonucleotide; multiple target; antisense treatment;  
 KW impaired respiration; inflammation; lung disease;  
 KW pulmonary vasoconstriction; inflammation; allergic rhinitis;  
 KW acute asthma; allergy; asthma; impeded respiration;  
 KW respiratory distress syndrome; pain; cystic fibrosis;  
 KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;  
 KW chronic obstructive pulmonary disease; leukemia; carcinoma;  
 KW colon cancer; breast cancer; lung cancer; pancreatic cancer;  
 KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;  
 KW prostate cancer; ss.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9213886-A1.  
 XX  
 PD 25-MAR-1999.  
 XX  
 PF 17-SEP-1998; 98WO-US19419.  
 XX  
 PR 09-JUN-1998; 98US-0093972.  
 PR 17-SEP-1997; 97US-0059160.  
 XX  
 PA (UWEC-) UNIV EAST CAROLINA.  
 XX  
 PI Nyce JW;  
 XX  
 DR WPI, 1999-229400/19.  
 XX  
 PT New antisense oligonucleotides used in treatment of, e.g. pulmonary  
 PT vasoconstriction  
 XX  
 PS Disclosure; Page 37; 120pp; English.  
 XX  
 CC The specification describes antisense oligonucleotides (AAX52869-X55271)  
 CC directed against at least 2 mRNAs selected from target genes, coding and  
 CC non-coding regions of RNAs corresponding to target genes, gene  
 CC initiation codons, genomic flanking regions, intron-exon borders, the  
 CC 5'-end, the 3'-end and the junction between coding and non-coding  
 CC regions and all segments of RNAs encoding proteins associated with one



DE Human neuronal NACHR alpha7 subunit encoding cDNA.  
 XX  
 KW Human; neuronal; nicotinic acetylcholine receptor; NACHR; drug screening;  
 XX Immunochemistry; NACHR alpha7 subunit; gene; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 73..1581  
 FT /tag= a  
 FT /product= "NACHR alpha7 subunit"  
 FT /note= "neuronal nicotinic acetylcholine receptor"  
 XX  
 PN WO200259266-A2.  
 XX  
 PD 01-AUG-2002.  
 XX  
 PF 29-OCT-2001; 2001MO-US50985.  
 XX  
 PR 01-NOV-2000; 2000US-0703951.  
 XX  
 TA (MERI ) MERCK & CO INC.  
 XX  
 CI Gillespie A, Claeps BO, Chavez-Noriega LE, Siegel R, Elliott KJ;  
 XX WPI; 2002-698532/75.  
 DR P-PSDB; AB882435.  
 XX  
 PT Cell comprising nucleic acids encoding human alpha and beta subunits of  
 PT neuronal nicotinic acetylcholine receptors, useful for in vitro  
 PT screening of a drug substance in a test system specific for humans  
 XX  
 PS Examples; Page 129-130; 143pp; English.  
 XX  
 CC The invention relates to a suitable host cell transfected with an  
 CC isolated nucleic acid molecule comprising a sequence of nucleotides or  
 CC ribonucleotides that encodes at least one alpha or beta subunit of a  
 CC human neuronal nicotinic acetylcholine receptor (NACHR). The compositions  
 CC and methods of the present invention, which provide a means to prepare  
 CC synthetic or recombinant receptors and receptor subunits that are  
 CC substantially free of contamination from many other receptor proteins,  
 CC are useful for observing the effect of a drug substance on a particular  
 CC subtype to perform in vitro screening of the drug substance in a test  
 CC system that is specific for humans. The antibodies can be used in  
 CC immunochemistry and for diagnostic and therapeutic applications. The  
 CC present sequence represents a human neuronal NACHR alpha7 subunit  
 CC encoding cDNA.  
 XX  
 SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 1 other;  
 XX  
 Query Match 27.8%; Score 127; DB 24; Length 1876;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-13;  
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 321 GGGCGAGGGCGAGGCCGGGCGACAGCCGAGACGTGAGCGCGCGCTGCTGCAGCT 380  
 DB 1 GGCCGAGGGCGAGGCCGGGCGACAGCCGAGACGTGAGCGCGCGCTGCTGCAGCT 60  
 QY 381 CCGGAGCTCAACATGCGCTGCTGCGGGAGGGCGTCTGGCTGGCGCGCTGCTGCTC 440  
 DB 61 CCGGAGCTCAACATGCGCTGCTGCGGGAGGGCGTCTGGCTGGCGCGCTGCTGCTC 120  
 QY 441 CTGCACG 447  
 DB 121 CTGCACG 127

RESULT 12  
 ID ABS54875  
 XX ABS54875 standard; cDNA; 1876 BP.  
 AC ABS54875;  
 XX

DT 06-DEC-2002 (first entry)  
 XX  
 DE Human neuronal nicotinic acetylcholine receptor alpha 7 subunit cDNA.  
 XX  
 KW Human; neuronal nicotinic acetylcholine receptor; nNACHR; gene; ss;  
 XX ion flux; alpha 7 subunit.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 73..1581  
 FT /tag= a  
 FT /product= "Human nNACHR alpha 7 subunit"  
 XX  
 PN US6440681-B1.  
 XX  
 PD 27-AUG-2002.  
 XX  
 PF 07-JUN-1995; 95US-0487596.  
 XX  
 PR 03-APR-1990; 90US-0504455.  
 XX 30-NOV-1992; 92US-0938154.  
 PR 08-MAR-1993; 93US-0028031.  
 PR 08-NOV-1993; 93US-0149503.  
 XX  
 TA (MERI ) MERCK & CO INC.  
 XX  
 CI Elliott KJ, Ellis SB, Harpold NM;  
 XX WPI; 2002-711528/77.  
 DR P-PSDB; AB670492.  
 XX  
 PT Identifying antagonists or agonists of human neuronal nicotinic  
 PT acetylcholine receptors, by contacting recombinant cells with test  
 PT compound, and measuring ion flux of cells or binding of compound to  
 PT nNACHR  
 XX  
 PS Claim 101; Column 57-60; 56pp; English.  
 XX  
 CC The invention relates to a method for identifying compounds that are  
 CC antagonists or agonists of human neuronal nicotinic acetylcholine  
 CC receptors (nNACHRs), by contacting recombinant cells with a test  
 CC compound and measuring ion flux, the electrophysiological response of the  
 CC cells or binding of the test compound to the nNACHR. The recombinant  
 CC cells are produced by transfection with a nucleic acid encoding at least  
 CC one human nNACHR (alpha or beta) subunit, such that the cells express an  
 CC nNACHR comprising one human subunit encoded by the transfected nucleic  
 CC acid. This sequence represents cDNA encoding the alpha 7 subunit of the  
 CC human nNACHR polypeptide.  
 XX  
 SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 1 other;  
 XX  
 Query Match 27.8%; Score 127; DB 24; Length 1876;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-13;  
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 321 GGGCGAGGGCGAGGCCGGGCGACAGCCGAGACGTGAGCGCGCGCTGCTGCAGCT 380  
 DB 1 GGCCGAGGGCGAGGCCGGGCGACAGCCGAGACGTGAGCGCGCGCTGCTGCAGCT 60  
 QY 381 CCGGAGCTCAACATGCGCTGCTGCGGGAGGGCGTCTGGCTGGCGCGCTGCTGCTC 440  
 DB 61 CCGGAGCTCAACATGCGCTGCTGCGGGAGGGCGTCTGGCTGGCGCGCTGCTGCTC 120  
 QY 441 CTGCACG 447  
 DB 121 CTGCACG 127

RESULT 13  
 ID AA051543/c  
 XX AA051543 standard; DNA; 3861 BP.  
 AC AA051543;  
 XX



AB013410-AB054121 represent genomic DNA sequences used to illustrate the method for determining the degree of cytosine methylation described in the disclosure of the invention.

Sequence 556 BP; 66 A; 71 C; 192 G; 227 T; 0 other;

Query Match 46.5%; Score 212.4; DB 24; Length 556;  
Best Local Similarity 68.4%; Pred. No. 2.5e-27;  
Matches 294; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

```

QY 28 CTTGGGAGGCGCCCTGCTGCGCCAGAGCGCGAGAGCGCGCTGCTGAGACT 87
DB 554 CTTTAAACAGCCCTTAACCTAACGAAACCGGAAACCGGAAACCGGCTCAATAAACT 495
QY 88 GGGGTGAGGTGCGCGGAGCGTACCGCGCGGAGTACTCTCCGCTCAACACTCGGG 147
DB 494 AAAAATAAAATATACCGGAAACGTAACCGCGGAAATATACCTCCGCTCAACCTCGAA 435
QY 148 CTGCACTTCCCTGGTGGCGCGCGGAGAGCTGCGCGGCTGAGAGGATGGCGGGCGGG 207
DB 434 CTACATTTCCCTAAATAATACCGCGGAAACGTAACCGCGGAAATATAAAATATAAAACGAA 375
/ 208 GACGAGGCGCGGCGCGGCTGCTCACTGAGAGAGCGCGCGGGCGGCGGGCGGG 267
DB 374 AACGAAACGAAACGAAACCTGCTACGTAATAAAACCGCGGAAACGAAACGAA 315
QY 268 GCGGCGCGCGGCTCTTAAAGCGCGGAGCGCGAGCGGAGTCTCTGTGCGCGCA 327
DB 314 ACGGCGCGCGGCTCTTAAAGCGCGGAGCGCGAGCGGAAATCTCTTAATACCGCA 255
QY 328 GCGGAGGCGCGGCGGAGCGCGAGAGCTGAGAGCGCGCGCTGCTGACCTCGGAGC 387
DB 254 AACCGAAACCGGAAACGAAACGTAATAACCGCGGAGCTGCTTAACCTCGGAAAC 195
QY 388 TCACATGCGCTGCTGCGCGGAGCGCTGCTGCGCGCTGCTGCTGAGAG 447
DB 194 TCACATGCGCTGCTGCGCGGAGCGCTGCTGCGCGCTGCTGCTGAGAG 135
QY 448 GTAAAGCCAC 457
DB 134 ATAAACACAC 125

```

RESULT 8  
AB047035

ID AB047035 standard; DNA; 556 BP.

AC AB047035;

XX 12-JUL-2002 (first entry)

Oligonucleotide for detecting cytosine methylation SEQ ID NO 33626.

Human: cytosine methylation; 5'-CPG-3'; uracil; cytosine; diagnosis;  
drug; side effect; cancer; central nervous system; cardiovascular;  
gastrointestinal; respiratory system; single nucleotide polymorphism;  
SNP; cell differentiation; ds.

OS Homo sapiens.

XX MO200218632-A2.

PN 07-MAR-2002.

XX 01-SEP-2001; 2001WO-EP10074.

XX 01-SEP-2000; 2000DE-1043826.

PR 05-SEP-2000; 2000DE-1044543.

XX (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K, Guetig D;

DR WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful  
PT for diagnosis and prognosis, comprises selective hybridization of  
PT amplicons from chemically treated DNA.

PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of  
CC methylation of a particular cytosine in a motif 5'-CPG-3', present in a  
CC genomic sample of DNA. The sample is treated chemically to convert  
CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
CC DNA that contains the target C is amplified to form a labeled amplicon.  
CC The amplicon is hybridized to two classes, each with at least one  
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers  
CC and the degree of hybridization to both classes is determined from the  
CC label on the amplicon. From the ratio of labels hybridized to the two  
CC classes of oligomers, the degree of methylation is calculated. The method  
CC is used: (i) for diagnosis and/or prognosis of side effects of  
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders  
CC of the central nervous, cardiovascular, gastrointestinal and respiratory  
CC systems etc., particularly by detecting mutations or single nucleotide  
CC polymorphisms (SNPs); and (ii) for differentiation of cell or tissue  
CC types and for investigating cell differentiation. The method allows the  
CC methylation status of many C residues to be determined simultaneously.  
CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the  
CC method for determining the degree of cytosine methylation described in  
CC the disclosure of the invention.

XX Sequence 556 BP; 227 A; 192 C; 71 G; 66 T; 0 other;

Query Match 46.5%; Score 212.4; DB 24; Length 556;  
Best Local Similarity 68.4%; Pred. No. 2.5e-27;  
Matches 294; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

```

QY 28 CTTGGGAGGCGCCCTGCTGCGCCAGAGCGCGAGCGCGCTGCTGAGACT 87
DB 3 CTTTAAACAGCCCTTAACCTAACGAAACCGGAAACCGGAAACCGGCTCAATAAACT 62
QY 88 GGGGTGAGGTGCGCGGAGCGTACCGCGCGGAGTACTCTCCGCTCAACACTCGGG 147
DB 63 AAAAATAAAATATACCGGAAACGTAACCGCGGAAATATAAAATATAAAACGAA 122
QY 148 CTGCACTTCCCTGGTGGCGCGCGGAGAGCTGCGCGGCTGAGAGGATGGCGGGCGGG 207
DB 123 CTACATTTCCCTAAATAATACCGCGGAAACCGTAACCGGAAATATAAAATATAAAACGAA 182
QY 208 GACGAGGCGCGGCGGCTGCTCACTGAGAGCGCGCGGCGCGGCGCGGGCGGG 267
DB 183 AACGAAACGAAACGAAACCTGCTCACTTAATAAAACCGCGGAAACGAAACGAA 242
QY 268 GCGGAGGCGCGGCGGCTTAAAGCGCGGAGCGCGGAGTACTCTGCTGCGGCA 327
DB 243 ACGCGGCGCGGACTCTTAAAGCGCGGAGCGGAAATCTCTTAACCGCA 302
QY 328 GCGGAGGCGCGGCGGAGCGGAGAGCTGAGAGCGCGGCTGCTGAGCTCGGAGC 387
DB 303 AACGAAACCGGAGAGCAACGAAACGTAATAACCGGAGTCTGCTCAACCTCGGAAAC 362
QY 388 TCACATGCGCTGCTGCGGAGAGCGTGTGCTGCGCTGCGCGGCTCTCTGACG 447
DB 363 TCACATGCGCTGCTGCGGAGAGCGTGTGCTGCGCTGCGCGGCTCTCTGACG 422
QY 448 GTAAAGCCAC 457
DB 423 ATAAACACAC 432

```

RESULT 9

AAV12197

ID AAV12197 standard; cDNA; 1876 BP.

XX AAV12197;

XX	ABJ34019/C
XX	ID ABJ34019 standard; DNA; 5304 BP.
XX	ABJ34019;
DT	26-MAR-2002 (first entry)
DE	Human immune system associated gene SEQ ID NO: 1992.
KW	Human, immune system disease; cytosine methylation; antiaethmatic; antiartherosclerotic; antihaemic; cytosatic; noctropic; neuroprotective; anti-HIV; anticomulant; ophthalmological; anthrumatic; arthritic; antidiabetic; antipsoriatic; antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia; acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy; neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene; ds.
OS	Homo sapiens.
PN	WO200200928-A2.
PD	03-JAN-2002.
PF	02-JUL-2001; 2001MO-EP07537.
PR	30-JUN-2000; 2000DF-1032529.
FR	01-SEP-2000; 2000DE-1045826.
PA	(EPIG-) EPIGENOMICS AG.
PI	Olek A, Piepenbrock C, Berlin K;
DR	WPI; 2002-130909/17.
XX	Nucleic acid comprising fragment of chemically modified gene, useful PT for diagnosis and treatment of diseases associated with abnormal PT cytosine methylation -
PS	Claim 1; SEQ ID NO 1992; 32pp + Sequence Listing; German.
CC	The present invention provides a number of human immune system associated CC genes which are modified by the methylation of cytosines. The sequences CC can be used in the diagnosis and treatment of immune system disorders, CC including eye diseases such as retinopathy, neovascular glaucoma and CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel CC diseases. The present sequence is a gene of the invention.
SO	Sequence 5304 BP; 1400 A; 152 C; 1205 G; 254 T; 0 other;
QY	Query Match 48.5%; Score 221.8; DB 24; Length 5304;
Bb	Beet Local Similarity 67.8%; Pred. No. 4.7e-29;
Matches	310; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
Db	1 AGAAGCGAAGGGAGGAGTGAAGCCTGGCTGGGAGGCCCTGGCTGGCCAGAGCGGC 60
QY	696 AAAACGCAAAAATAAAATTAACCTTAACCTTAACAACCCCTTAACCAAAAAACGG 637
Db	61 AGGCGGAGAGCCGCTCGGTGAGACTGGGGGTGAGGTGCCCCGAGAGCTACCCAGCGCC 120
QY	636 AAACCGAAAAACCCGCTGATTAATAAACCTAAATAATAATTAATCCGAAAGCTACCCAGCC 577
Db	121 GGAGGTACTCTCCGCTCACACCTGGGGTGCAATTTCCTGGGTGGCCCGCGAGAGCTGG 180
QY	576 GAAATAACTCTCCGCTCACACCTGAACTCAATTCCTTAATAATAACCGCGAAAGCTTA 517
Db	181 CCCGGGCTGGAGGAGTGGCGGGGCGGAGCGGGGCGGGGGCTGTCACGTGGAG 240
QY	516 CCCGAACTAAAAAATTAATAAAACGAAACGAAACGAAACGAAACCTGTCAGTAA 457

QY	241	AGCGCGCGGGGGCGGGCGGGGGGGCGCGCGCGCGCTCTTAAAGCGCGCAGCC	300
Db	456	AAACGCGGAAAACGAAACGAAAACGAAAACGCGCGCCACCTCTTAAAAACGCGCAAC	397
QY	301	GAGCGCGAGGCTGCCTCTTCTGTGGCGCGCAGCGCGAGGCCCGGAGCAGCAGCGTGGAG	360
Db	396	GAAAGAGAAATACCTCTATTAACCGCAAAACGCAAAACCGAAGCAGAACGAAACGTAATA	337
QY	361	CGCGCGCGCTCGCTGCGAGCTCCGGAGCTCAACATGCGTCTGCGCGGAGCGCTTGGC	420
Db	336	CGCGCGCAGCTGCTACCACTCCGAAACCTCAACATACGCTACCTCGCGAAAAACGTTAAC	277
QY	421	TGGCGCTGCGCGCGCTGCTCCGACGGAAGAGCAGC	457
Db	276	TAAAGCTAACCGCTGCTGCTCTTACAGATTAACAC	240
RESULT 7			
ABQ47034/c			
ID	ABQ47034	standard; DNA; 556 BP.	
AC	ABQ47034;		
XX			
DT	12-JUL-2002	(first entry)	
XX			
DE	Oligonucleotide for detecting cytosine methylation SEQ ID NO 33625.		
XX			
KW	Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;		
KW	drug; side effect; cancer; central nervous system; cardiovascular;		
KW	gastrointestinal; respiratory system; single nucleotide polymorphism;		
KW	SNP; cell differentiation; ds.		
OS			
XX	Homo sapiens.		
XX			
XX	WO200218632-A2.		
PN			
PD	07-MAR-2002.		
XX			
PF	01-SEP-2001; 2001WO-EP10074.		
XX			
XX	01-SEP-2000; 2000DE-1043826.		
PR	05-SEP-2000; 2000DE-1044543.		
XX			
XX	(EPIG-) EPIGENOMICS AG.		
PA			
XX			
PI	Olek A, Piepenbrock C, Berlin K, Guetig D;		
DR	WPI; 2002-371829/40.		
XX			
PT	Determining the degree of cytosine methylation in genomic DNA, useful		
PT	for diagnosis and prognosis, comprises selective hybridization of		
PT	amplicons from chemically treated DNA.		
XX			
XX	Claim 12; 56pp + Sequence Listing: 56pp; German.		
XX			
CC	This invention describes a novel method for determining the degree of		
CC	methylation of a particular cytosine in a motif 5'-CpG-3', present in a		
CC	genomic sample of DNA. The sample is treated chemically to convert		
CC	cytosine (C) but not methylated C, to uracil, then part of the genomic		
CC	DNA that contains the target C is amplified to form a labeled amplicon.		
CC	The amplicon is hybridised to two classes, each with at least one		
CC	member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers		
CC	and the degree of hybridisation to both classes is determined from the		
CC	label on the amplicon. From the ratio of labels hybridised to the two		
CC	classes of oligomers, the degree of methylation is calculated. The method		
CC	is used: (i) for diagnosis and/or prognosis of side effects of		
CC	therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders		
CC	of the central nervous, cardiovascular, gastrointestinal and respiratory		
CC	systems etc., particularly by detecting mutations or single nucleotide		
CC	polymorphisms (SNP's), and (ii) for differentiation of cell or tissue		
CC	types and for investigating cell differentiation. The method allows the		
CC	methylation status of many C residues to be determined simultaneously.		

therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders of the central nervous, cardiovascular, gastrointestinal and respiratory systems etc., particularly by detecting mutations or single nucleotide polymorphisms (SNP's) and (ii) for differentiation of cell or tissue types and for investigating cell differentiation. The method allows the methylation status of many C residues to be determined simultaneously. ABO1310-AB054121 represent genomic DNA sequences used to illustrate the method for determining the degree of cytosine methylation described in the disclosure of the invention.

SQ Sequence 556 BP; 57 A; 71 C; 241 G; 187 T; 0 other;

Query Match	Score	DB	Length
65.3%	298.4	24	556
81.1%	139.41		

Oy	26	GGGCTTGGGACAGCCCTGGGCTGGGCGAGAGGCGAGAGCCGAGAGCCGCTCGGTGAGA	85
Db	1	GGTTTGGGTAGTTTTGGTTGGTTGGTTAGAGCGCGAGGTCGAGAGTTGTTCCGTGGAGA	60
Oy	86	CTGGGGGTGAGAGTCCCGGAGCGTACCCAGCGCCGGAGTACTCCCGCTACACTCG	145
3	61	TTGGGGGTGAGAGTGTTCGGAGCGTATTAGGTCGGAGATTTTTCGTTATATTTCG	120
Oy	146	GGCTGAGATTCCTGGGTGGCGCGGAGACGCTGGCCCGGGCTGAGAGATGGCGGGCG	205
Db	121	GGTTGTAGTTTTTGGGTGCTGTCGAGACGTTGGTTCCGGTTGAGAGGATGGGAGGCG	180
Oy	206	GGGACGGGGCGGGGGCGGGGCTGTCACTGTGAGAGCGCGCGGGGCGGGCGG	265
Db	181	GGGACGGGGCGGGGGCGGGGCTGTCTAGTGTGAGAGCGCGCGGGGGCGGGCGG	240
Oy	266	GGGCGCGCCGCCGCTCTTAAAGCGCGCCAGCCGACGCGGAGATGCTCTGTGGCCG	325
Db	241	GGCGCGCGCTTCGTTTTTAAAGCGCGCCAGTGGAGCGCGGAGGTGTTTTGTGTGCG	300
Oy	326	CAGGCGAGGCCCGGGCGACAGCCGAGACGTGAGCGGGCGGCGCTGTCAGCTCCGGG	385
Db	301	TAGGCGTAGGTTGGGGCGATGTCGAGACGTGAGACGCTGCGTTCTGTGTAGTTTCGG	360
Oy	386	ACTCAACATGCGCTGCTCGCCGGAGGCGCTGGCTGGCGCTGACCGCGTGCCTCTGCA	445
Db	361	ATTTAATATCGTTGTTTCGTCGGGAGGCGTTTGGTTGGCGTTGGTGCCTGTTTGTGA	420
Oy	446	CGGTAAAG	453
Db	421	CGGTAAAG	428

ABQ47037 standard; DNA; 556 BP.

DT 12-JUL-2002 (first entry)

DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 33628.

KM Human, cytosine methylation; 5'-bp-g-3'; uracil; cytosine; diagnosis;  
KM drug; side effect; cancer; central nervous system; cardiovascular;  
KM gastrointestinal; respiratory system; single nucleotide polymorphism  
KM SNP; cell differentiation; ds.

OS Homo sapiens.

PN WO200218632-A2

PD 07-MAR-2002

PF 01-SEP-2001; 2001WO-EP10074

PR 01-SEP-2000; 2000DE-1043826

PR 05-SEP-2000; 2000DE-1044543

PA (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K, Guetig D;

DR WPI; 2002-371829/40

**PT** Determining the degree of cytosine methylation in genomic DNA, useful for diagnosis and prognosis. comprises selective hybridization of

PT amplicons from chemically treated DNA  
xy

PS Claim 12; 56pp + Sequence Listing; 56pp; German  
YX

This invention describes a novel method for determining the degree of methylation of a particular cytosine in a motif 5'-CpG-3', present in a genomic sample of DNA. The sample is treated chemically to convert cytosine (C) but not methylated C, to uracil, then part of the genomic DNA that contains the target C is amplified to form a labeled amplicon. The amplicon is hybridised to two classes, each with at least one member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the degree of hybridisation to both classes is determined from the label on the amplicon. From the ratio of labels hybridised to the two classes of oligomers, the degree of methylation is calculated. The method is used: (i) for diagnosis and/or prognosis of side effects of therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders of the central nervous, cardiovascular, gastrointestinal and respiratory systems etc., particularly by detecting mutations or single nucleotide polymorphisms (SNP's); and (ii) for differentiation of cell or tissue types and for investigating cell differentiation. The method allows the methylation status of many C residues to be determined simultaneously. ABO1310-AB054121 represent genomic DNA sequences used to illustrate the method for determining the degree of cytosine methylation described in the disclosure of the invention.

556 BP; 187 A; 241 C; 71 G; 57 T; 0 other.

Query Match	65.3%;	Score 298.4;	DB 24;	Length 556;
-------------	--------	--------------	--------	-------------

QY	26	GGCCCTGGGGCAGCCCTGGACCTGGCCGAGAGGCGCGAGGCGCAGAGACCCGCTCGGTGGAGAGA	85
Db	556	GGTTTGGGTAGTCTTTTGGTTGGTTGGTTAGAGCGCCAGAGTCCAGAGTTCGTTGGTGGAGA	497
QY	86	CTGGGGGTGAGAGTCCCGGAGCGTACCAGCGCCGGAGTACCTCCGCTCACACTCG	145
Db	496	TTGGGGGGTGAAGTCTTTCGAGACGTAATTACGCTGGAGTATTTTTCGTTATTTTCG	437
QY	146	GGCTGCAATTCCTCGGTGGGCGCGCGAGAACGCTGGCCCGGGCTGGAGAGGAATGGCGGGGCG	205
Db	436	GGTTGTATTTTTTTGGGTGTGTCGGAACGTTGGTTTGGGGTTGAGAGGATATGGGGGGCG	377
QY	206	GGGACGCGGGGCGGGGCGGGGCTGCTCACTGAGAGGCGCGCGGGGCGGGGCGGGGCGG	265
Db	376	GGAGACGGGGGGGGGGCGGGGTCGTTACGTGAGAGGCGCGCGGGGCGGGGCGGGGCGG	317
QY	266	GGGCGCGCGCCCGCTCTTTAAAGCGCGCGAGCGCGAGCGCGAGAGTGCCTCTGTGGCG	325
Db	316	GGGCGCGCGGTTGCGTTTTTAAAGCGCGCGAGCTGAGACGGCGGAGAGTGTTTTGTGGTCG	257
QY	326	CAGGCGCAGGCGCGGCGGAGACCGAGACGTGAGACGCGCGGCTGCGTCACTCCGG	385
Db	226	TAGGCGTAGGTTCCGGGCATTACTCGAAGCGTGAAGCGCGCTCGTTTCGTTGAGTTCCGG	197
QY	386	ACTCAACTGCGCTGCTGCGCGGAGAGGCGTCTGCGTGCCTGCGCTGCGCTCCTGCA	445
Db	196	ATTTAATATGCGTTGTTGTCGGGAGGCGTTTGGTTGCGGTTGTCGCGCTCCTTTTGTGA	137
QY	446	CGGTAAG 453	
Db	136	CGGTAAG 129	



```

Db      301 GAGCGCGGAGGTGCTCTGTGGCCGAGCGGAGCGCCGGGGGAGACAGCGGAGCGTGGAG 360
QY      361 GCGCGCGGCTCGCTGCAAGTCCGGGAGCTCAAC 392
      |||||
Db      361 GCGCGCGGCTCGCTGCAAGTCCGGGAGCTCAAC 392
      |||||

RESULT 3
ID      ABLJ34018 standard, DNA, 5304 BP.
XX
XX      ABLJ34018;
XX
XX      26-MAR-2002 (first entry)
XX
XX      Human immune system associated gene SEQ ID NO: 1991.
XX
XX      Human immune system associated gene; cytosine methylation; antiasthmatic;
XX      antiarteriosclerotic; antianaemic; cytosarctic; noctropic;
XX      neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX      antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
XX      antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX      acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX      neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
XX      gene; ds.
XX
XX      Homo sapiens.
XX
XX      OS
XX      PN
XX      FN
XX      WO200200928-A2.
XX
XX      03-JAN-2002.
XX
XX      02-JUL-2001; 2001WO-EP07537.
XX
XX      30-JUN-2000; 2000DE-1032529.
XX      PR
XX      01-SEP-2000; 2000DE-1043826.
XX
XX      (EPIG-) EPIGENOMICS AG.
XX
XX      Olek A, Piepenbrock C, Berlin K;
XX
XX      WPI: 2002-130909/17.
XX
XX      Nucleic acid comprising fragment of chemically modified gene, useful
XX      for diagnosis and treatment of diseases associated with abnormal
XX      cytosine methylation -
XX
XX      Claim 1; SEQ ID NO 1991, 32pp + Sequence Listing; German.
XX
XX      The present invention provides a number of human immune system associated
XX      CC genes which are modified by the methylation of cytosines. The sequences
XX      CC genes which are modified by the methylation of cytosines. The sequences
XX      CC can be used in the diagnosis and treatment of immune system disorders,
XX      CC including eye diseases such as retinopathy, neovascular glaucoma and
XX      CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX      CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX      CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX      CC diseases. The present sequence is a gene of the invention.
XX
XX      Sequence 5304 BP; 1480 A; 152 C; 1219 G; 2453 T; 0 other;
XX
XX      Query Match      69.7%; Score 318.6; DB 24; Length 5304;
XX      Best Local Similarity 81.5%; Pred. No. 3.2e-45;
XX      Matches 369; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
XX
QY      1 AGAAGCGAAGGAGGAGGTGAGAGCCGCTGCGGAGAGCCCGCTGCGGAGAGGAGCGG 60
      |||||
Db      4609 AGAAGCGTAAAGGAGAGGTGAGAGTTGTTTGGGTAGTTTGTGTTTGAAGGCGCG 4666
      |||||
QY      61 AGGCGGAGAGCCCGCTCGGTGAGAGCTGGGGGTGAGAGTGCCTCGGAGAGTACCCAGCGCC 120
      |||||
Db      4669 AGGTTCGAGAGTTCGTTGCTGTGGAGATTGGGGGTGAGAGTGTTCGAGAGCTATTAGCGCT 4722
      |||||
QY      121 GGGAGTACTCTCCCGCTACACCTCGGCTGCAGATTCCCTGTGGTGGCGCGGAGAGCGCTGG 180
      |||||

```

Dd	4729	GGAGCTATTTTCTTTATATTTTCGGTTGATTTTTTTGGATGCTGTCGACGTTGG	4788
Qy	181	CCCCGGCTGAGGAGATGCGGGGCGGGAGACGGGGCGGGGCGGCTGTCACGTGAG	240
Dd	4789	TTCCGGTTGAGAGGATGGGGGGGCGGGAGCGGGGCGGGGCTTCGTACGTGGAG	4848
Qy	241	AGCGCGCGGGGGCGCGGGCGGGGCGCGCGCTCTTAAAGCGCGGAGCC	300
Dd	4849	AGCGCGCGGGGGCGGGGCGGGGCGGGGCGGCTTCGTTCGTTTAAAGCGCGGAGTC	4908
Qy	301	GAGCGGGAGAGTGCCTCTGTGCGCCGACGGCGGAGCCCGGCGACACCGAGACTGGAG	360
Dd	4909	GAGCGGGAGAGTGTTTTGTGTGTGTAAGGGGTAGTTCGGGCGATATCGAGAGTGGAG	4968
Qy	361	CGCGCGGCTCGCTGCAGCTCCGGGACTCAACATGCGCTGTCGCGGGAGAGCTGTGAC	420
Dd	4969	CGCGCTCGGTTCTGTGTAGTTTCGGGATTTAATATGCTTTGCTCGGAGAGGCTTTGGT	5028
Qy	421	TGGCGCTGGCGCGCTGCTCCTCGACCGTAAG	453
Dd	5029	TGGCGTTGTGTCGGCTGTTTTCGACGTAAG	5061
RESULT 4			
ID	ABQ47036	standard; DNA; 556 BP.	
AC	ABQ47036;		
DX	12-JUL-2002	(first entry)	
XX			
DE		Oligonucleotide for detecting cytosine methylation SEQ ID NO 33627.	
XX			
KM	Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;		
KW	drug; side effect; cancer; central nervous system; cardiovascular;		
KW	gastrointestinal; respiratory system; single nucleotide polymorphism;		
XX	SNP; cell differentiation; ds.		
OS	Homo sapiens.		
XX			
XX	WO200218632-A2.		
PN			
PD	07-MAR-2002.		
XX			
PF	01-SEP-2001; 2001MO-EP10074.		
XX			
XX	01-SEP-2000; 2000DE-1043826.		
PR			
PR	05-SEP-2000; 2000DE-1044543.		
XX			
XX	(EPIC-) EPIGENOMICS AG.		
PA			
XX	Olek A, Piepenbrock C, Berlin K, Gnetig D;		
PI			
XX	WPI; 2002-371829/40.		
DR			
PT	Determining the degree of cytosine methylation in genomic DNA, useful		
PT	for diagnosis and prognosis, comprises selective hybridization of		
PT	amplicons from chemically treated DNA		
XX			
PS	Claim 12; 56bp + Sequence Listing; 56pp; German.		
XX			
CC	This invention describes a novel method for determining the degree of		
CC	methylation of a particular cytosine in a motif 5'-CpG-3', present in a		
CC	genomic sample of DNA. The sample is treated chemically to convert		
CC	cytosine (C) but not methylated C, to uracil, then part of the genomic		
CC	DNA that contains the target C is amplified to form a labeled amplicon.		
CC	The amplicon is hybridised to two classes, each with at least one		
CC	member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers		
CC	and the degree of hybridisation to both classes is determined from the		
CC	label on the amplicon. From the ratio of labels hybridised to the two		
CC	classes of oligomers, the degree of methylation is calculated. The method		
CC	is used: (1) for diagnosis and/or prognosis of side effects of		



DR WPI, 1999-288306/24.  
 XX Human alpha-7 neuronal nicotinic acetylcholine receptor and related  
 PT polynucleotides  
 XX Claim 1, Fig 4, 104pp; English.  
 XX The present invention describes an isolated nucleotide sequence (1)  
 CC encoding at least a portion of the human alpha-7 neuronal nicotinic  
 CC acetylcholine receptor (alpha7-hnAChR). Also described are: (1) a  
 CC peptide encoded by (1); (2) a vector comprising (1); (3) a host cell  
 CC transformed with a vector of (2); (4) a polynucleotide comprising at  
 CC least 15 nucleotides which hybridises under stringent conditions to at  
 CC least a portion of (1); (5) a method for detection of a polynucleotide  
 CC encoding alpha 7-hnAChR in a biological sample; and (6) a method for  
 CC amplification of nucleic acid from a sample suspected of containing  
 CC nucleic acid encoding alpha 7-hnAChR. The primers and probes from the  
 CC present invention can be used on brain tissue and blood samples of  
 CC humans suspected of suffering from schizophrenia, small cell lung  
 CC carcinoma, breast cancer and nicotine-dependent illness. This is  
 CC particularly useful for diagnosis of schizophrenia. Other illnesses  
 CC that can be studied/diagnosed are epilepsy (e.g. juvenile myoclonic  
 CC epilepsy) and Prader-Willi and Angelman's syndromes.  
 .X Sequence 457 BP; 58 A; 145 C; 204 G; 50 T; 0 other;  
 SQ  
 Query Match 100.0%; Score 457; DB 20; Length 457;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-68;  
 Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGAAGCGAAGGAGAGAGTAAAGCTTGGGCGACCCCTGGGCTGGCCAGAGGGCGG 60  
 Db 1 AGAAGCGAAGGAGAGAGTAAAGCTTGGGCGACCCCTGGGCTGGCCAGAGGGCGG 60  
 QY 61 AGGCGGAGAGCCGCTCGTGGAGACATCGGGGCTGAGAGTGGCCGAGACCGAGCGCC 120  
 Db 61 AGGCGGAGAGCCGCTCGTGGAGACATCGGGGCTGAGAGTGGCCGAGACCGAGCGCC 120  
 QY 121 GGGAGTACCTCCGCTCACAACCTCGGCTGAGATTCCCTGAGTGGCCGCGAGACGCTGG 180  
 Db 121 GGGAGTACCTCCGCTCACAACCTCGGCTGAGATTCCCTGAGTGGCCGCGAGACGCTGG 180  
 QY 181 CCCGGGCTGAGAGGATGCGGGGCGGGGAGAGGGGGCGGGGCTCTGTCACGTGGAG 240  
 Db 181 CCCGGGCTGAGAGGATGCGGGGCGGGGAGAGGGGGCGGGGCTCTGTCACGTGGAG 240  
 QY 241 AGGCGCGAGGGGCGGGGCGGGGCGGGGCGGGGCTCTTAAAGCGGGCGAGCGC 300  
 Db 241 AGGCGCGAGGGGCGGGGCGGGGCGGGGCGGGGCTCTTAAAGCGGGCGAGCGC 300  
 QY 301 GAGCGGCGAGGTGCTCTGTGTGGCGGAGCGCAGAGCCCGGGCGACCGAGACGTGGAG 360  
 Db 301 GAGCGGCGAGGTGCTCTGTGTGGCGGAGCGCAGAGCCCGGGCGACCGAGACGTGGAG 360  
 QY 361 CGGCGCGGCTGCGTGGAGGCTCCGGGACTCAATGCGCTGCTGCGGGAGGGGTCTGGC 420  
 Db 361 CGGCGCGGCTGCGTGGAGGCTCCGGGACTCAATGCGCTGCTGCGGGAGGGGTCTGGC 420  
 QY 421 TGGCGCTGAGCGGCTGCTCTGTCACCGTAAAGCCAG 457  
 Db 421 TGGCGCTGAGCGGCTGCTCTGTCACCGTAAAGCCAG 457

RESULT 2  
 AAX56151  
 ID AAX56151 standard; DNA; 392 BP.  
 XX  
 AC AAX56151;  
 XX  
 DT 15-JUL-1999 (first entry)  
 XX  
 DE Human alpha-7 nicotinic receptor nucleotide sequence fragment #8.  
 XX

KW Human: alpha-7 nicotinic receptor; neuronal; hybridisation; probe;  
 KW alpha-7 neuronal nicotinic acetylcholine receptor; schizophrenia;  
 KW small cell lung carcinoma; breast cancer; nicotine-dependent illness;  
 KW epilepsy; juvenile myoclonic epilepsy; Prader-Willi syndrome;  
 KW Angelman's syndrome; PCR primer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO920757-A2.  
 XX  
 PD 29-APR-1999.  
 XX  
 PF 15-OCT-1998; 98MO-US21762.  
 XX  
 PR 23-OCT-1997; 97US-0956518.  
 XX  
 PA (FREE/) FREEDMAN R.  
 PA (LEON/) LEONARD S.  
 XX  
 PI Freedman R, Leonard S;  
 XX  
 DR WPI, 1999-288306/24.  
 XX  
 PT Human alpha-7 neuronal nicotinic acetylcholine receptor and related  
 PT polynucleotides  
 XX  
 PS Claim 1, Fig 8; 104pp; English.  
 XX  
 CC The present invention describes an isolated nucleotide sequence (1)  
 CC encoding at least a portion of the human alpha-7 neuronal nicotinic  
 CC acetylcholine receptor (alpha7-hnAChR). Also described are: (1) a  
 CC peptide encoded by (1); (2) a vector comprising (1); (3) a host cell  
 CC transformed with a vector of (2); (4) a polynucleotide comprising at  
 CC least 15 nucleotides which hybridises under stringent conditions to at  
 CC least a portion of (1); (5) a method for detection of a polynucleotide  
 CC encoding alpha 7-hnAChR in a biological sample; and (6) a method for  
 CC amplification of nucleic acid from a sample suspected of containing  
 CC nucleic acid encoding alpha 7-hnAChR. The primers and probes from the  
 CC present invention can be used on brain tissue and blood samples of  
 CC humans suspected of suffering from schizophrenia, small cell lung  
 CC carcinoma, breast cancer and nicotine-dependent illness. This is  
 CC particularly useful for diagnosis of schizophrenia. Other illnesses  
 CC that can be studied/diagnosed are epilepsy (e.g. juvenile myoclonic  
 CC epilepsy) and Prader-Willi and Angelman's syndromes.  
 CC  
 SQ Sequence 392 BP; 51 A; 122 C; 180 G; 39 T; 0 other;  
 Query Match 85.8%; Score 392; DB 20; Length 392;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-57;  
 Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGAAGCGAAGGAGAGTAAAGCTTGGGCGACCCCTGGGCTGGCCAGAGGGCGG 60  
 Db 1 AGAAGCGAAGGAGAGTAAAGCTTGGGCGACCCCTGGGCTGGCCAGAGGGCGG 60  
 QY 61 AGGCGGAGAGCCGCTCGTGGAGACATCGGGGCTGAGAGTGGCCGAGACCGAGCGCC 120  
 Db 61 AGGCGGAGAGCCGCTCGTGGAGACATCGGGGCTGAGAGTGGCCGAGACCGAGCGCC 120  
 QY 121 GGGAGTACCTCCGCTCACAACCTCGGCTGAGATTCCCTGAGTGGCCGCGAGACGCTGG 180  
 Db 121 GGGAGTACCTCCGCTCACAACCTCGGCTGAGATTCCCTGAGTGGCCGCGAGACGCTGG 180  
 QY 181 CCCGGGCTGAGAGGATGCGGGGCGGGGAGAGGGGGCGGGGCTCTGTCACGTGGAG 240  
 Db 181 CCCGGGCTGAGAGGATGCGGGGCGGGGAGAGGGGGCGGGGCTCTGTCACGTGGAG 240  
 QY 241 AGGCGCGGCGGGGCGGGGCGGGGCGGGGCGGGGCTCTTAAAGCGCGGAGCC 300  
 Db 241 AGGCGCGGCGGGGCGGGGCGGGGCGGGGCGGGGCTCTTAAAGCGCGGAGCC 300  
 QY 301 GAGCGGCGAGGTGCTCTGTGTGGCGGAGCGCAGAGCCCGGGCGACCGAGACGTGGAG 360  
 Db 301 GAGCGGCGAGGTGCTCTGTGTGGCGGAGCGCAGAGCCCGGGCGACCGAGACGTGGAG 360

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2003, 05:13:39 ; Search time 269 Seconds  
(without alignments)  
4586.034 Million cell updates/sec

Title: US-08-956-518a-94

Perfect score: 457

Sequence: 1 AGAAGCGAAGGAGAGGTAG.....CTCCTGCACGGTAAGCCAC 457

Scoring table:  
IDENTITY NUC  
Gapop 10.0, Gapext 1.0

rched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

N Geneseq 190Jun03.\*  
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	457	100.0	457	20	AA56144
2	392	85.8	392	20	AA56151
3	318.6	69.7	5304	24	ABU34018
4	298.4	65.3	556	24	ABQ47036
5	298.4	65.3	556	24	ABQ47037
6	221.8	48.5	5304	24	ABU34019
7	212.4	46.5	556	24	ABQ47034
8	212.4	46.5	556	24	ABQ47035

Result No.	Score	Query Match	Length	ID	Description
9	127	27.8	1876	15	AAV12197
10	127	27.8	1876	18	AA748239
11	127	27.8	1876	24	ABV73248
12	127	27.8	1876	24	AB54875
13	69.2	15.1	3861	14	AAQ51543
14	68.8	15.1	114955	20	AA53491
15	67	14.7	600	24	ABQ52496
16	67	14.7	600	24	ABQ52497
17	66.2	14.5	1000	21	AA02484
18	65.6	14.4	1127	21	AA02477
19	65.6	14.4	114793	22	AA08215
20	65.2	14.3	1337	20	AA217263
21	65.2	14.3	114955	20	AA53491
22	64.6	14.1	12733	24	ABK98631
23	64.6	14.1	12739	24	ABK98592
24	64	14.0	3198	20	AAK02974
25	63.2	13.8	43058	24	ABN97455
26	63.2	13.8	43058	24	ABL64982
27	63.2	13.8	43058	24	ABL65219
28	63	13.8	154746	24	AA025519
29	63	13.8	154746	24	AA025519
30	62	13.6	1964	24	ABZ11298
31	61.6	13.5	544	24	ABQ24355
32	61.6	13.5	544	24	ABQ24355
33	61.6	13.5	22976	20	AA583426
34	61.6	13.5	41936	24	ABL67924
35	61.4	13.4	2307	24	ABK32842
36	61.2	13.4	1416	25	ABZ20967
37	60.6	13.3	320	21	AA38185
38	60.6	13.3	1281	13	AA032896
39	60.2	13.2	712	24	ABQ40858
40	60.2	13.2	712	24	ABQ40859
41	59.8	13.1	434	22	AA184343
42	59.8	13.1	687	24	ABQ19226
43	59.8	13.1	687	24	ABQ19227
44	59.8	13.1	754	24	ABQ40800
45	59.8	13.1	754	24	ABQ40801

#### ALIGNMENTS

RESULT 1  
ID AAX56144 standard; DNA: 457 BP.  
AC AAX56144;  
XX  
XX  
DT 15-JUL-1999 (first entry)  
XX  
XX Human alpha-7 nicotinic receptor nucleotide sequence fragment #1.  
DE  
XX  
XX Human: alpha-7 nicotinic receptor; neuronal; hybridisation; probe:  
KW alpha-7 neuronal nicotinic acetylcholine receptor; schizophrenia;  
KW small cell lung carcinoma; breast cancer; nicotine-dependent illness;  
KW epilepsy; juvenile myoclonic epilepsy; Prader-Willi syndrome;  
KW Angelman's syndrome; PCR primer; ss.  
XX  
OS Homo sapiens.  
XX  
XX  
XX WO920757-A2.  
XX  
XX PD 29-APR-1999.  
XX  
XX PF 15-OCT-1998; 98WO-US21762.  
XX  
XX PR 23-OCT-1997; 97US-0956518.  
XX  
XX (FREE/) FREEDMAN R.  
XX (LEON/) LEONARD S.  
XX  
XX Freedman R, Leonard S;  
XX

LOCATION: (745)..(747)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (750)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (752)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (754)..(756)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (761)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (763)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (818)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (843)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (892)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (897)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (920)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (941)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (993)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (995)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1066)..(1067)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1082)..(1084)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1105)..(1107)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1130)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1147)

OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1162)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1202)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1233)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1247)..(1248)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1261)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1485)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1500)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1544)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1601)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1605)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1635)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1637)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1654)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1674)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1695)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1852)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1866)..(1965)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: (1968)..(1980)  
OTHER INFORMATION: a, t, c, g, unknown or other

```

: SEQUENCE DESCRIPTION: SEQ ID NO: 11:
: US-10-349-836-11
:
: Query Match 27.8%; Score 127; DB 12; Length 1876;
: Best Local Similarity 100.0%; Pred. No. 1.8e-19;
: Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0
:
: QY 321 GCGCGAGCGCCAGAGCCCGGGCGAGACGCGAGAGCGCGCGCTCGCTGAGCT 380
: Db 1 GCGCGAGCGCCAGAGCCCGGGCGAGACGCGAGAGCGCGCGCTCGCTGAGCT 60
:
: QY 381 CCGGAGCTCAACATGCGTGTCTCGCGGAGAGCGCTGTGCGCTGGCCGCTGCTC 440
: Db 61 CCGGAGCTCAACATGCGTGTCTCGCGGAGAGCGCTGTGCGCTGGCCGCTGCTC 120
:
: QY 441 CTGCACG 447
: Db 121 CTGCACG 127
:
: RESULT 5
: US-10-017-161-1857
: Sequence 1857, Application US/10017161
: Publication No. US20030143668A1
: GENERAL INFORMATION:
: APPLICANT: SUMA, MAKIKO
: APPLICANT: ASAI, KIYOSHI
: APPLICANT: AKIYAMA, YUTAKA
: APPLICANT: ABURATANI, HIROYUKI
: TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
: FILE REFERENCE: 084335/0152
: CURRENT APPLICATION NUMBER: US/10/017,161
: CURRENT FILING DATE: 2002-12-18
: PRIOR APPLICATION NUMBER: JP 2001/246789
: PRIOR FILING DATE: 2001-06-18
: NUMBER OF SEQ ID NOS: 2430
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1857
: LENGTH: 3163
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: source
: LOCATION: (1)..(3163)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (201)..(2963)
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (556)..(655)
: OTHER INFORMATION: a, t, c, g, unknown or other
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (658)..(669)
: OTHER INFORMATION: a, t, c, g, unknown or other
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (671)..(697)
: OTHER INFORMATION: a, t, c, g, unknown or other
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (704)..(707)
: OTHER INFORMATION: a, t, c, g, unknown or other
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (709)..(736)
: OTHER INFORMATION: a, t, c, g, unknown or other
: FEATURE:
: NAME/KEY: modified_base
: LOCATION: (739)..(743)
: OTHER INFORMATION: a, t, c, g, unknown or other
: FEATURE:
: NAME/KEY: modified_base

```



Query				Description	
No.	Score	Match Length	DB	ID	
1	318.6	69.7	5304	12	US-10-311-455-1991
2	221.6	48.5	5304	12	Sequence 1991, A
3	127	27.8	1876	10	US-09-311-455-1992
4	127	27.8	1876	12	Sequence 7, Appl
5	70.6	15.4	3163	12	US-10-349-836-11
6	67.2	14.7	3133	12	US-10-017-161-1857
7	65.6	14.4	114793	12	Sequence 1857, A
8	65	14.2	956	13	US-10-148-806-3
9	64.6	14.1	12733	14	Sequence 1483, A
10	64.6	14.1	12739	14	Sequence 3, Appl
11	64	14.0	152331	13	Sequence 31508, A
12	63.2	13.8	43058	10	Sequence 47, Appl
13	63.2	13.8	43058	10	Sequence 8, Appl
14	63.2	13.8	43058	10	Sequence 16, Appl
15	62.8	13.7	1117	12	Sequence 292, Appl
16	61.6	13.5	41936	10	Sequence 529, Appl
					Sequence 1305, Appl
					Sequence 1490, Appl
					Sequence 1163, Appl

Db  
4669 AGTCGAGAGTTCGTCGTCGAGATTGGG

17	61.4	13.3	4	2307	11	US-09-693-5198-87	Sequence 87, Appl
18	59.4	13.0	5	5452	12	US-10-017-161-1481	Sequence 1481, Appl
19	58	12.7	2561	10	US-09-976-740-48	Sequence 48, Appl	
20	58	12.7	2561	13	US-10-023-523-48	Sequence 48, Appl	
21	58	12.7	2561	13	US-10-023-523-48	Sequence 48, Appl	
22	57.8	12.6	922	12	US-10-017-161-1435	Sequence 1435, Appl	
23	56.8	12.4	692	13	US-10-027-632-96253	Sequence 96253, Appl	
24	56.8	12.4	692	13	US-10-027-632-96253	Sequence 96253, Appl	
25	55	12.0	2678	12	US-10-017-161-1647	Sequence 1647, Appl	
26	54.8	12.0	10144	10	US-09-880-107-2168	Sequence 2168, Appl	
27	54.4	11.9	1614	10	US-09-976-740-45	Sequence 45, Appl	
28	54.4	11.9	1614	10	US-10-023-523-45	Sequence 45, Appl	
29	54.4	11.9	1614	13	US-10-023-523-45	Sequence 45, Appl	
30	54.4	11.9	12425	10	US-09-976-740-50	Sequence 50, Appl	
31	54.4	11.9	12425	13	US-10-023-523-50	Sequence 50, Appl	
32	54.4	11.9	12425	13	US-10-023-523-50	Sequence 50, Appl	
33	54.4	11.9	33023	10	US-09-880-107-73350	Sequence 3350, Appl	
34	54.2	11.9	22005	12	US-10-292-198-2	Sequence 2	
35	54	11.8	594	12	US-10-140-472-10	Sequence 10, Appl	
36	54	11.8	594	12	US-10-141-761-10	Sequence 10, Appl	
37	54	11.8	594	12	US-10-142-885-10	Sequence 10, Appl	
38	54	11.8	594	14	US-10-123-155-10	Sequence 10, Appl	
39	54	11.8	594	15	US-10-146-731-10	Sequence 10, Appl	
40	54	11.8	883	13	US-10-027-632-4358	Sequence 4358, Appl	
41	54	11.8	2000	10	US-09-901-488A-180	Sequence 180, Appl	
42	54	11.8	2000	10	US-09-853-526-180	Sequence 180, Appl	
43	54	11.8	56516	10	US-09-901-488A-1	Sequence 1, Appl	
44	54	11.8	56516	10	US-09-853-526-1	Sequence 1, Appl	
45	54	11.8	56520	10	US-09-901-488A-179	Sequence 179, Appl	

	Query Match	Best Local Similarity	69.7%;	Score 318.6;	DB 12;	Length 5304;
	Matches 369;	Conservative	0;	Mismatches 84;	Indels 0;	Gaps 0;
QY	1	AGAACCCAGAGGAGAGCTGTAAGCCTGGCCCTTTGGGCAACCCCTGGCCCTGGCCAGAGGCGCC	60			
Db	4609	AGAACCTTAAAGGAGAGGAGTGAAGTTGGTTTGGGTACTTTTGGTTTGGTTTGAAGGCGCCG	4668			
QY	61	AGGCCGAGAGCCCGCTGGGTGAGACTGAGGAGTGAGGTCGCCGAGCGTATCCAGCGCC	120			
Db	4669	AGGTCGAGAGTTTGGTTTCGGTGAGATTTGGGGGTGGAGAGTGTTCGAGACGTAATTAGGCTC	4728			

Seq primer: RM1 TAGACTCACTATAGGAGA

Class: BAC ends

High quality sequence start: 25

High quality sequence stop: 160.

FEATURES  
Location/Qualifiers

1. 1198

/organism="Gallus gallus"

/mol\_type="genomic DNA"

/strain="Red Jungle Fowl"

/db\_xref="taxon:9031"

/clone="CH261-138F22"

/sex="female"

/cell\_line="UCD001, inbred 256"

/clone\_id="CH261"

/note="Vector: pTRABAC2.1; Site 1: EcoRI; Site 2: EcoRI; CH261 Female Chicken library - For library and clone ordering information: http://www.chori.org/bacpac"

BASE COUNT 29 a 748 c 341 g 35 t 45 others

ORIGIN

Query Match 19.0%; Score 87; DB 29; Length 1198;

Best Local Similarity 49.2%; Pred. No. 9.7e-06;

Matches 207; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

QY 14 GAGTAAAGCTTGCCCTTGGGACACCCCTTGCCCTGCGACAGGCGCGAGCCGAGACCC 73

Db 675 GGG 616

QY 74 GCTCGTGAAGACTGAGGAGTGGAGTGGCCCGGAGCGTACCAGCGCGGGAGTACTCC 133

Db 615 GCGGG 556

QY 134 GCTCACAACCTCGGAGCTCAGTTCCTTGGTGGTGGCCCGAGACGCTGCGCCGCGTGGAG 193

Db 555 GGG 496

QY 194 GATGCGGG 253

Db 495 GCGGG 436

QY 254 CGGG 313

Db 435 CCGGG 376

QY 314 CTTCTGTGGCGGACGAGCGGCGGCGGACGACGAGTGAAGCGCGCGCGCTGCGC 373

Db 375 GCGCGGG 316

QY 374 TGCACTCCGGGACTCAATGCGCTCTGCGCGGAGGCGTCTGCGCTGCGCTGCGCGC 433

Db 315 GCGGG 256

QY 434 G 434

Db 255 G 255

RESULT 15  
CNS04CNH 549 bp DNA linear GSS 01-SEP-2000LOCUS  
DEFINITION  
09J04 of library G from Tetraodon nigroviridis, genomic surveyACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISMAL284678  
GI:8023072  
GSS; genome survey sequence.  
Tetraodon nigroviridis  
Tetraodon nigroviridis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodontidae; Tetraodon.REFERENCE  
1

Search completed: September 7, 2003, 15:14:42

Job time : 2272 secs

AUTHORS  
Roest Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,  
Bernot, A., Pizames, C., Wincker, P., Brothier, P., Querier, F.,  
Saurin, W., and Weissbach, J.  
Estimate of human gene number provided by genome-wide analysis  
using Tetraodon nigroviridis DNA sequence  
Nat. Genet. 25 (2), 235-238 (2000)TITLE  
JOURNAL  
MEDLINE  
PUBMED  
20296633  
10835645REFERENCE  
AUTHORS  
Roest Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,  
Bernot, A., Pizames, C., Wincker, P., Brothier, P., Querier, F.,  
Saurin, W., Bernot, A., and Weissbach, J.  
Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis  
Genome Res. 10 (7), 939-949 (2000)TITLE  
JOURNAL  
MEDLINE  
PUBMED  
20359837  
10899143REFERENCE  
AUTHORS  
Direct Submission  
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetraodon nigroviridis  
genome. For more information, please take a look at  
http://www.genoscope.cns.fr/Tetraodon.FEATURES  
source  
1. 549  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone="09J04"  
/clone\_id="G"  
/note="Genoscope sequence ID : COBG09DE02LP1-end : T7"

BASE COUNT 20 a 81 c 407 g 16 t 25 others

ORIGIN  
Query Match 18.9%; Score 86.2; DB 29; Length 549;  
Best Local Similarity 48.4%; Pred. No. 1.4e-05;  
Matches 196; Conservative 13; Mismatches 196; Indels 0; Gaps 0;

QY 10 GGGAGAGTAAAGCTTGGGAGTGGGAGCCCTTGGGAGTGGGAGCCCGAGAGCCGAGAGA 69

Db 120 GGG 179

QY 70 GCGCGTGGTGAAGCTGAGGAGTGAAGTGGCCGAGCGTACCGAGCGCGGAGTACC 129

Db 180 GCGCGCGCGGG 239

QY 130 TCCGCTCAACCTCGGCGTGCATTTCTTGGGTGGCGCGGAGACGCTGACCCGGGCTG 189

Db 240 GSGCGCGGG 239

QY 190 GAGGAGTGGCGGG 249

Db 300 CGGG 359

QY 250 GGG 369

Db 360 GGG 419

QY 310 GGTGCTCTGTGGCGGAGCGGAGCGCGGCGGAGCGGAGCGGAGCGGAGCGGCGCGC 369

Db 420 GGG 479

QY 370 TGCTGAGCTCCGGGACTCAATGCGCTCTGCGCGGAGCGC 414

Db 480 TSGGGGAGACGCGATGTGCGCTGCTGCTGAGAGGGGAGCG 524



```

DEFINITION      603181651F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5245812 5',
ACCESSION      B1917899
VERSION        B1917899.1
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 651)
AUTHORS       NIH-MGC http://mgc.nci.nih.gov/.
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished.
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgabbs-remail.nih.gov
               Tissue Procurement: Life Technologies, Inc.
               cDNA Library Preparation: Life Technologies, Inc.
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LNL at:
               http://image.llnl.gov
               Plate: LLM11620 row: e column: 13
               High quality sequence stop: 647.
               Location/Qualifiers
                 1..651
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /clone="IMAGE:5245812"
                   /lab_host="DH10B"
                   /clone_11b="NIH_MGC_121"
                   /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
                   Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
                   fetal brains, female age 20 weeks, female age 24 weeks,
                   and male age 26 weeks. Library is oligo-dT primed and
                   directionally cloned (EcoRV site is destroyed upon
                   cloning). Average insert size 1.7 kb, insert size range
                   0.7-3.5 kb. Library is normalized and enriched for
                   full-length clones and was constructed by C. Gruber
                   (Invitrogen). Research Genetics tracking code 017. Note
                   this is a NIH_MGC Library."
BASE COUNT      137 a 188 c 182 g 144 t
ORIGIN
Query Match    19.3%; Score 88; DB 12; Length 651;
Best Local Similarity 100.0%; Pred. No. 6.8e-06;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY              354 CGTGGAGCGCGCCGGCTGCTGCAAGCTCCGGAGACTCAACATGCGCTGCGCGGAGAGC 413
                |||
                1 CGTGGAGCGCGCCGGCTGCTGCAAGCTCCGGAGACTCAACATGCGCTGCTCGCGGAGAGC 60
DY              414 GTCTGGCTGGCGGCTGGCGCGCGCTGCTCC 441
                |||
                61 GTCTGGCTGGCGGCTGGCGCGCGCTGCTCC 88
LOCUS           BX403654 1103 bp mRNA linear EST 15-MAY-2003
LOCUS           BX403654 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBAA0072H12.1
ACCESSION      BX403654
VERSION        BX403654
KEYWORDS       3-PRIME, mRNA sequence.
SOURCE         EST.
ORGANISM       Homo sapiens (human)
               Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 1103)
AUTHORS       Li, W.B., Gruber, C., Jeessee, J., and Polayes, D.
TITLE         Full-length cDNA libraries and normalization

```

[illegible]



ACCESSION	BX415111
VERSION	BX415111.1
KEYWORDS	GI:30765470
SOURCE	EST.
ORGANISM	Homo sapiens (human)
JOURNAL	Homoe sapiens
COMMENT	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 982) Li,W.B., Gruber,C., Jesse,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished Contact : Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0CAP004BP10QPL. Location/Qualifiers .982 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0CAP004YI20" /tissue_type="THYMUS" /clone_id="Homo sapiens THYMUS" /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
BASE COUNT	177 a 257 c 323 g 18 t 207 others
ORIGIN	
Query Match	20.3%; Score 92.6; DB 13; Length 982;
Best Local Similarity	32.4%; Pred.No.1.le-06;
Matches 143; Conservative 117; Mismatches 179; Indels 2; Gaps 1;	
Oy	1 AGAAGCGAAGGAGGTAGAGCCTGGGCTTGAGGAGCCGCTGCAGAGGCGCG 60
Dd	105 AAAAAARRGGGGGGGGGGGGSSSGVGVSSSSSSSSSSSSCCCCYGgg 164
Oy	61 AGCGCGAGAGCCCGCTCGGTGAGACTGGGGATGAGGTGCCCGAGAGCTGACCAGCGCC 120
Dd	165 GGGGCGGSVSGSgCCCCSSCGKCCCCKCCCCCGGGGGGGCGCCCCSSSSCCCC 224
Oy	121 GGGAATCCTCCCGCTCAACACTTGGGCTGCAATTCCCTGGGTGGCCCGCAGACGCTGG 180
Dd	225 SCSSGSCCCCCCCCCCCCCCCCCCCCCCGCCGCCCGCGGAGCCTCCSSCGSSSS 284
Oy	181 CCCGGGCTGAGGGAGTGGCGGGGGGAGCGGGGGCGGGGCGTCGTCACGTGAG 240
Dd	285 CSSSSGCSGSGSSCGSSGGGGGGGSSSCSSGSSGGGSSSSSG--SGSCCGGGGGG 342
Oy	241 AGCGCGCGGGGCGGGGCGGGGCGGGGCGCGCGCCGCTCTTAAGGCGCGCGAGCC 300
Dd	343 SSCSSSGCGSGCGCGGGGGSSSCGSSCGSSSGGGGGSSSGGGGSSSGGGGMS 402
Oy	301 GAGCGCGAGGTGCTCTGTGTGCGCGCAGGCGCAGGCCCGGCGACAGCCGAGAGCTGAG 360
Dd	403 SSSSGGCCCCCGCCSSCGGGGSSGGS CGSSGSGSVGGGSSCGGCGSSCGSGSGGGG 462
Oy	361 CGCGCGGCTCGCTGCGACACTCGGGAGTCACATGCGCTGTCCCGGGAGAAGCGTGGC 420
Dd	463 GGSccsssggggcccccgaggscggggggggcscsbgggggggggggsgsgscgcggc 522
Oy	421 TGGGCTGGCGCGCTGCTCC 441
Dd	523 CCGMGGGGCCCGCGCCSSCC 543

CSNS015Y4	LOCUS	1203 bp	DNA	linear	GSS 26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC				
DEFINITION	BACN15E10 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL106054				
VERSION	AL106054.1				
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peiryoeta; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1203) Genoscope. Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
AUTHORS	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (BDGP) - http://www.iedgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.				
COMMENT	Location/Qualifiers				
FEATURES	1. 1203 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACN15E10" /clone_1lb="DrosBAC" /plasmid="pBelobAC11" /note="end : T7"				
BASE COUNT	154 a 274 c 380 g 158 t 237 others				
ORIGIN					
Query Match	19.9% Score 91; DB 29; Length 1203;				
Best Local Similarity	38.1%; Pred. No. 2e-06;				
Matches 164; Conservative	84; Mismatches 160; Indels 2; Gaps 1;				
30	TTGGGCAAGCCCTGCTGCTGCAGAGGCGGAGCCGAGACCCGCTGCTGAGACTG 89				
Db	673 TTCCGGGGSSSSSSSSSSCCSCGCCGCCSSGGGGGGCGCGSSGGGGGGGGGG 732				
90	GGGTGAGAGTCCCGGAGCGTACCCAGGCGCGGAGTACTCCCGCTACACCTTGGGCT 149				
Db	733 GGG 792				
150	GCAATTCCTGG 207				
793	GG 852				
208	GACGG 267				
853	CCGG 912				
268	GCGCGCCGCCCGGCTCTTTAAAGCGCGCGAGCCGAGCGGCGAGGTGCTTGTGGCCGA 327				
Db	913 GNGGGGGGGGMMVGGSGGGVVGGGVMMVGGSGGMMVGGMAVGGGGGMMSCMGGVSSSS 972				
328	GGGGCGAGGG 387				
973	VGGSSSSGG 1032				
388	TCAACATGCGCTGCTCCCGGAGGCGCTGTGCTGCGCTGCGCGCTGCTGCGAGC 447				
Db	1033 GMSGGGGSSSSGG 1092				
448	GTAAAGCCAC 457				



5'-GACTAGTTTATGATCGGAGCGCCGCTT(15)-3'. Size selected > 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."

BASE COUNT 198 a 251 c 247 g 212 t

Query Match 27.1%; Score 124; DB 13; Length 908;  
Best Local Similarity 100.0%; Pred. No. 5e-12;  
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 324 CGAGGCGCAGGCGCGGCGCAGACCGAGAGCTGAGCGCGGCTGCTGAGCTCCG 383  
DB 1 CGAGGCGCAGGCGCGGCGCAGACCGAGAGCTGAGCGCGGCTGCTGAGCTCCG 60  
QY 384 GGACTCAACATGCGCTGCTGCGCGGAGGCGCTGCTGCGCTGCTGCTCTG 443  
61 GGACTCAACATGCGCTGCTGCGCGGAGGCGCTGCTGCGCTGCTGCTCTG 120  
QY 444 CAGC 447  
DB 121 CAGC 124

RESULT 4  
LOCUS B0070018 1080 bp mRNA linear EST 02-APR-2002

DEFINITION AGENCOURT\_6660390 NIH\_MGC\_47 Homo sapiens cDNA clone IMAGE:5922336

ACCESSION B0070018 5', mRNA sequence.

VERSION B0070018.1 GI:19899064

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov

Plate: LCM2089 row: b column: 01  
High quality sequence stop: 142.  
Location/Qualifiers  
1. 1080  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5922336"  
/tissue\_type="neuroblastoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: brain; Vector: pORF7; Site1: XhoI; Site2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 259 a 286 c 244 g 290 t 1 others

Query Match 24.9%; Score 113.8; DB 13; Length 1080;  
Best Local Similarity 94.4%; Pred. No. 2.7e-10;  
Matches 118; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 322 GCGCAGGCGCAGGCGCGGCGCAGACCGAGAGCTGAGCGCGGCTGCTGAGCTC 381  
DB 1 GCGCAGGCGCAGGCGCGGCGCAGACCGAGAGCTGAGCGCGGCTGCTGAGCTC 60  
QY 382 GCGGACTCAACATGCGCTGCTGCGCGGAGGCGCTGCGGCTGCGGCTGCTGCTC 441  
DB 61 GCGGACTCAACATGCTGCTGCTGCGCGGAGGAGCTGCGGCTGCGGCTGCTGCTC 120  
QY 442 TGCAC 446  
DB 121 TGCAC 125

RESULT 5  
LOCUS BX403124 1034 bp mRNA linear EST 13-MAY-2003

DEFINITION BX403124 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens

ACCESSION BX403124 cDNA clone CS00C0041A12 5-PRIME, mRNA sequence.

VERSION BX403124.1 GI:30611171

KEYWORDS EST.

SOURCE Homo sapiens (human);

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: sequef@genoscope.cns.fr. Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 7646.r For

more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS1DC001ZG090P1&cluster=7646.r. Contact :  
Feng Liang Email : fliang@life.techn.com URL :  
http://fulllength.invitrogen.com/invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS1DC001ZG090P1.

Location/Qualifiers  
1. 1034  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC0041A12"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 213 a 298 c 288 g 218 t 17 others

Query Match 22.3%; Score 101.8; DB 13; Length 1034;  
Best Local Similarity 84.2%; Pred. No. 3e-08;  
Matches 123; Conservative 1; Mismatches 21; Indels 1; Gaps 1;

QY 302 ACGGCGAGTCTCTGTCGCGGAGGCGCGGCGCAGACCGAGAGCTGAGCTC 361  
DB 43 ACGGCGAGTCTCTGTCGCGGAGGCGCGGCGCAGACCGAGAGCTGAGCTC 102  
QY 362 GCGCGGCTGCTGTCGAGCTCGGAGACTCAACATGCGCTGCTGCGGAGGCGCTGCTGCT 421  
DB 103 GCGCGGCTGCTGTCGAGCTCGGAGACTCAACATGCGCTGCTGCGGAGGCTGCTGCTGCT 161

## High quality sequence stop: 404.

FEATURES  
Location/Qualifiers

1. 1125  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6338722"  
/tissue\_type="neuroblastoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_47"  
/note="Organ: Brain; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."  
Note: this is a NIH\_MGC Library."

BASE COUNT  
192 a 482 c 242 g 209 t

Query Match  
Best Local Similarity 100.0%; Pred. No. 3.2e-13;  
Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 CTGTGCGCGAGCGCGAGCGCGCGAGCGCGAGCGCGCGCGCTGCTGC 376  
Db 71 CTGTGCGCGAGCGCGAGCGCGCGAGCGCGAGCGCGCGCGCTGCTGC 130  
QY 377 AGCTCCGGAGCTCAACATGCTGCTGCTGCGGAGCGCTGCTGCGCTGCGCCGCTC 436  
Db 131 AGCTCCGGAGCTCAACATGCTGCTGCTGCGGAGCGCTGCTGCGCTGCGCCGCTC 190  
QY 437 GCTCTGTCAGC 447  
Db 191 GCTCTGTCAGC 201

RESULT 2  
LOCUS BQ894356 1084 bp mRNA linear EST 16-AUG-2002  
DEFINITION AGENCOURT\_8725920 NIH\_MGC\_47 Homo sapiens cDNA clone IMAGE:6338722  
5', mRNA sequence.

ACCESSION BQ894356  
VERSION BQ894356.1 GI:22286370  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1084)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LNCM2532 row: 9 column: 11  
High quality sequence stop: 482.

FEATURES  
Location/Qualifiers

1. 1084  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6338722"  
/tissue\_type="neuroblastoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_47"  
/note="Organ: Brain; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."  
Note: this is a NIH\_MGC Library."

BASE COUNT  
232 a 335 c 287 g 229 t 1 others

Query Match  
Best Local Similarity 100.0%; Pred. No. 6.9e-13;  
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 GTGGCGCGAGCGCGAGCGCGCGAGCGCGAGCGCGCGCGCTGCTGCAG 378  
Db 1 GTGGCGCGAGCGCGAGCGCGCGAGCGCGAGCGCGCGCGCTGCTGCAG 60  
QY 379 CTCCGGAGCTCAACATGCTGCTGCTGCGGAGCGCTGCTGCGCTGCGCGCTGC 438  
Db 61 CTCCGGAGCTCAACATGCTGCTGCTGCGGAGCGCTGCTGCGCTGCGCGCTGC 120  
QY 439 TCCTGTCAGC 447  
Db 121 TCCTGTCAGC 129

RESULT 3  
LOCUS BU149265 908 bp mRNA linear EST 03-SEP-2002  
DEFINITION AGENCOURT\_8443358 Lupski sympathetic\_trunk Homo sapiens cDNA clone IMAGE:6192234 5', mRNA sequence.

ACCESSION BU149265  
VERSION BU149265.1 GI:22662797  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 908)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. James R. Lupski  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLMW13593 row: 0 column: 19  
High quality sequence stop: 598.

FEATURES  
Location/Qualifiers

1. 908  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6192234"  
/sex="male"  
/tissue\_type="sympathetic trunk"  
/dev\_stage="adult, 16 yr"  
/lab\_host="DH10B"  
/clone\_lib="Lupski\_sympathetic\_trunk"  
/note="Vector: pCMV-SPORT6 (Life Technologies); Site: 1: NotI; Site: 2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TGCACCAAGCGCTCG-3' and

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2003, 05:33:14 ; Search time 2268 Seconds  
(without alignments)  
4897.331 Million cell updates/sec

Title: US-08-956-518A-94

Sequence: 1 AGAAGCGACGAGAGAGTAG.....CTCTGACAGGTAAAGCCAC 457

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST.\*  
1: em\_estda.\*  
2: em\_esthm.\*  
3: em\_estin.\*  
4: em\_estnu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hic.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hic.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pin.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_tod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*  
29: gb\_gss2.\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	131	28.7	1125	13	BUI48759
2	129	28.4	1084	13	B0894356
3	124	27.1	908	13	BUI49265
4	113.8	24.9	1080	13	B0070018

5	101.8	22.3	1034	13	BX403124
6	96.2	21.1	222	13	BX374871
7	93	20.4	1201	13	BX365176
8	92.6	20.3	982	13	BX415111
9	91	19.9	1203	29	CNS01514
10	90.2	19.7	1201	9	AL530299
11	88.6	19.4	1023	29	AG128304
12	88	19.3	651	12	B1917899
13	87	19.0	1103	13	BX403654
14	87	19.0	1198	29	CC190008
15	86.2	18.9	549	29	CNS04041
16	85	18.6	888	12	BG809572
17	85	18.6	888	13	BX496688
18	85	18.5	1057	13	BX496688
19	85	18.6	1364	12	BM810045
20	84.8	18.6	807	13	BX464554
21	84.6	18.5	924	13	BX442207
22	84.4	18.5	761	28	A2184492
23	84.4	18.5	943	12	BG785567
24	83.8	18.3	1201	9	AL563250
25	83.6	18.3	1050	13	B0716109
26	83.4	18.2	1065	13	B0681076
27	83	18.2	811	12	B1956687
28	82.8	18.1	932	29	CNS00720
29	82.6	18.1	1026	13	CNS01523
30	82.4	18.0	1101	29	CNS01523
31	82.4	18.0	1201	13	BX335523
32	82.2	18.0	1080	28	AO893056
33	82	17.9	440	13	BX403888
34	81.8	17.9	776	29	CNS01091
35	81.8	17.9	800	12	BG780621
36	81.8	17.9	888	29	AG030591
37	81.8	17.9	1061	13	BX391246
38	81.6	17.9	1003	29	AG159111
39	81.4	17.8	836	28	A2185303
40	81.4	17.8	857	29	AG061653
41	81.2	17.8	827	28	A2183795
42	81.2	17.8	925	29	CNS0091P
43	81.2	17.8	942	29	AG120142
44	81	17.7	932	29	CNS0072Q
45	80.8	17.7	1158	29	CC251504

## ALIGNMENTS

RESULT 1  
BUI48759  
LOCUS  
DEFINITION BUI48759 1125 bp mRNA linear EST 03-SEP-2002  
ACCESSION AGENCOURT 8695252 NIH\_MGC\_47 Homo sapiens cDNA clone IMAGE:6377441  
VERSION BUI48759.1 GI:22662291  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)\*

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Homo sapiens (human)\*  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 1125)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Straubeberg, Ph.D.  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
<http://image.llnl.gov>  
Plate: LLCM2561 row: d column: 18  
High quality sequence start: 54



Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
 Series: IRAX Plate: 69 Row: C Column: 18  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21536283.

## FEATURES

## Source

1.1810  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:45359 IMAGE:5245812"  
 /tissue\_type="Brain, fetal, whole pooled"  
 /clone\_id="NIH\_MGC\_121"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 393.1358  
 /codon\_start=1  
 /product="unknown (protein for MGC:45359)"  
 /protein\_id="AAH37571.1"  
 /db\_xref="GI:22902223"  
 /translation="MOBADISGYIPNGEMDLVGIPIGRSERFYECKEPPDYFTVT  
 MRRTLYGMLIPCVLISALILVFLPADSGEKISLIGTVLLSLTFMLLVABIM  
 PATSDSVPLIAQVPASTIIVGLSVVTVTVLYOHHHPDGGKPKRTVILLNCAIM  
 FLNRKRGEEDKVRACQHKQRCSLASVEMSAVAPPASNGNLLYIGFRGLDGVHCVP  
 TPDSGVVCGRNACSPFDEHLLHGGPPEDPDIAKILEEVRYIANRFRCDSEAVC  
 SEWKFACVVDRLCLMAFSVFTICTIGILMSAPNFEAVSKDFA"

## CDS

BASE COUNT 380 a 530 c 495 g 405 t  
 ORIGIN

## Query Match

Best Local Similarity 20.6%; Score 94; DB 9; Length 1810;  
 Matches 94; Conservative 0; Pred. No. 0.00019; Mismatches 0; Indels 0; Gaps 0;

QY	354	CGTGGAGCGCGCGGCTGCTGTCAGACTCCGGGACTCAACATGCGCTGCTGCGCGGAGGC	413
Db	1	CGTGGAGCGCGCGGCTGCTGTCAGACTCCGGGACTCAACATGCGCTGCTGCGCGGAGGC	60
QY	414	GTCGTGGCTGGCGCTGGCCGCGCTGCTCTGTCACG	447
Db	61	GTCGTGGCTGGCGCTGGCCGCGCTGCTCTGTCACG	94

Search completed: September 7, 2003, 14:36:43  
 Job time : 2504 secs

/translation="MRCSPGVWVLAALASLHVSLOGEFORLYKEVKNVPLERPV  
 ANDSOPLTVFSLQIMDVDEKNOVLTNTNLMOWMTDHLQMNVSSEYRQVTRF  
 PDGQIKWPDILLNSADREDFATHTNTVNSGHCQYLPPEIFKSSCTIDVRFEPD  
 VOCHLKFGSNGYGGMSLDLQWQADISGTYINGEMDLVGIKRSERIECKSEYP  
 DVTFVTMRRTLLVYGNLIPCVLISALVFLIPADSGEKISLIGITVLSLTFM  
 LLVAEIMPATSDSVPLIAQVPASTWIVGLSVVTVIVLOYHHDPDGGKMPKTRVI  
 LLNMCAMFLMRKRGEDKVRPAOKORCSLASVEMSAVAPPPASNGNLTYIGRGL  
 DGVHCVPDSDGSGVCGRMACSPTHDEHLHGOGPEGGPDIAKIIIEBRYIANRRQ  
 DESAVGSEMKRPAACVDRCLMAFVSFTIITIGILMSAPNFAVASKDFA"  
 sig\_peptide 73..138  
 mat\_peptide 139..1578  
 /product="nicotinic acetylcholine receptor alpha7 subunit"  
 1582..1876  
 BASE COUNT 369 a 553 c 531 g 423 t  
 ORIGIN  
 Query Match 27.8%; Score 127; DB 9; Length 1876;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-09;  
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 321 GGCCGAGGCGGCGGCGGAGACGCTGAGAGCCGCGGCTCGCTCAGCT 380  
 DB 1 GGCCGAGGCGGCGGCGGAGACGCTGAGAGCCGCGGCTCGCTCAGCT 60  
 QY 381 CGGAGCTCAACATCGCTGCTCGCGGAGAGCGCTGAGAGCCGCGGCTC 440  
 DB 61 CGGAGCTCAACATCGCTGCTCGCGGAGAGCGCTGAGAGCCGCGGCTC 120  
 QY 441 CTGCACG 447  
 DB 121 CTGCACG 127

RESULT 14  
 HSA7A7A 2087 bp mRNA linear PRI 01-JUN-1994  
 LOCUS H.sapiens mRNA for neuronal nicotinic acetylcholine receptor  
 ACCESSION X70297.1 GI:496606  
 KEYWORDS neuronal nicotinic acetylcholine receptor alpha-7 subunit.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 Peng, X., Katz, M., Gerzanich, V., Anand, R., and Lindstrom, J.  
 Human alpha 7 acetylcholine receptor: cloning of the alpha 7  
 subunit from the SH-SY5Y cell line and determination of  
 pharmacological properties of native receptors and functional alpha  
 7 homomers expressed in Xenopus oocytes  
 JOURNAL Mol. Pharmacol. 45 (3), 546-554 (1994)  
 MEDLINE 94195283  
 PUBMED 8145738  
 2 (bases 1 to 2087)  
 REFERENCE Katz, M.  
 DIRECT SUBMISSION  
 Submitted (04-FEB-1993) M. Katz, Univ of Pennsylvania School of  
 Medicine, Dept of Neuroscience, 36th & Hamilton Walk, Rm 235  
 Stemmler Hall, Philadelphia, PA 19104, USA  
 JOURNAL Location/Qualifiers  
 1. 2087  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="SHSY3"  
 /cell\_line="SHSY-5Y"  
 /cell\_type="neuroblastoma"  
 /tissue\_type="brain"  
 /clone\_id="lambda zap II"  
 104..1612  
 /codon\_start=1  
 /product="neuronal nicotinic acetylcholine receptor"

FEATURES  
 source

CDS

alpha-7 subunit"  
 .//protein\_id="CAA49778.1"  
 .//db\_xref="GI:496607"  
 /db\_xref="SWISS-PROT:P36544"  
 /translation="MRCSPGVWVLAALASLHVSLOGEFORLYKEVKNVPLERPV  
 ANDSOPLTVFSLQIMDVDEKNOVLTNTNLMOWMTDHLQMNVSSEYRQVTRF  
 PDGQIKWPDILLNSADREDFATHTNTVNSGHCQYLPPEIFKSSCTIDVRFEPD  
 VOCHLKFGSNGYGGMSLDLQWQADISGTYINGEMDLVGIKRSERIECKSEYP  
 DVTFVTMRRTLLVYGNLIPCVLISALVFLIPADSGEKISLIGITVLSLTFM  
 LLVAEIMPATSDSVPLIAQVPASTWIVGLSVVTVIVLOYHHDPDGGKMPKTRVI  
 LLNMCAMFLMRKRGEDKVRPAOKORCSLASVEMSAVAPPPASNGNLTYIGRGL  
 DGVHCVPDSDGSGVCGRMACSPTHDEHLHGOGPEGGPDIAKIIIEBRYIANRRQ  
 DESAVGSEMKRPAACVDRCLMAFVSFTIITIGILMSAPNFAVASKDFA"  
 sig\_peptide 104..175  
 mat\_peptide 176..1609  
 /product="unnamed"  
 BASE COUNT 441 a 601 c 573 g 472 t  
 ORIGIN  
 Query Match 26.6%; Score 121.6; DB 9; Length 2087;  
 Best Local Similarity 96.1%; Pred. No. 2.4e-08;  
 Matches 146; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 296 GAGCCGAGCGGCGGAGTCCCTCTGTGCGGAGCGAGCCCGGCGACAGCGA 355  
 DB 9 GAGCCGAGCGGCGGAGTCCCTCTGTGCGGAGCGAGCCCGGCGACAGCGA 66  
 QY 356 TGGAGCGGCGGCGGCGGAGTCCGCGGAGCTCAACATGGCGTCTCGCGGAGGCGT 415  
 DB 67 TGGAGCGGCGGCGGCGGAGTCCGCGGAGCTCAACATGGCGTCTCGCGGAGGCGT 126  
 QY 416 CTGCTGCGGCGGCGGCGGAGTCCGCGGAGCTCAACATGGCGTCTCGCGGAGGCGT 447  
 DB 127 CTGCTGCGGCGGCGGCGGAGTCCGCGGAGCTCAACATGGCGTCTCGCGGAGGCGT 158

RESULT 15  
 BC037571 1810 bp mRNA linear PRI 16-SEP-2002  
 LOCUS Homo sapiens, clone MGC:45359 IMAGE:5245812, mRNA, complete cds.  
 DEFINITION BC037571  
 ACCESSION BC037571  
 VERSION BC037571.1 GI:22902222  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1810)  
 AUTHORS Strausberg, R.  
 JOURNAL DIRECT SUBMISSION  
 Submitted (13-SEP-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
 Akher, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,  
 Maduro, O.L., Maslell, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Stancirip, S., Thomas, P.J., Touchman, J.W.,  
 Tsougen, C., Vogt, J.D., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.

```

translation="MRCSPGCVWALAAALHLSVLOSQFORLKYELVYONPLERF  

ANDSPLITVEFSILIOIMDYDERKNOVLFTNMLMSMTDHLQMNSEYEGVATRF  

PDIOTMKPDLILNSADERDATPFHTNVLYVNSSGCOTLPRIPISSCTIDYRFPDP  

VOHKLRFGSMYSYGSLDLOQWEADISGIYENGEMDLVGIPKRSEFYECCKEPTF  

LVETVKFRRTTLYGYNLIDPEVALISALALLFLPADBSGEKSLGITVLSTLMFM  

LLVMEIAPSDSPILAQYFASTMIVGLSVTVTVIYLQYHHHPDDGKKPKMTRVYI  

LLNMCAWFLEMRKEPGEKRYPAACOHKORCSLASVEMSAYAPPASNNLLIYGFREG  

DGVCVTPPSGSVQCVRMCSCSFPHDSHLLHGSGPREGPDLAKILEERYIANFPROO  

DESAFVSSEKFFAACVDRLCMAFSVFITICTIGILMSAPNFVAWSKDA"  

1582..1876  

a 553 c 531 g 423 t  

27.8%; Score 127; DB 6; Length 1876;  

Identity 100.0%; Pred.No.4.2e-09;  

Conservative 0; Mismatches 0; Indels 0; Gaps 0;  

AGGGCCAGGCCCGGGGCGCAGCCGAGACGTGAGACGGCGCGCTGCTGCTGCT 380  

AAGCCACAGCCCGGGGCGCAGCCGAGACGTGAGACGGCGCGCTGCTGCTGCT 60  

CTCAACAATGCGCTGCTGCTGCGCGGAGAGCGTGCTGGCTGGCGTCGCGCTC 440  

CTCAACAATGCGCTGCTGCTGCGCGGAGAGCGTGCTGGCTGGCGTCGCGCTC 120  

G 447  

|  

G 127  

cortic acetylcholine receptor alpha7 subunit precursor,  

complete cds.  

1876 bp mRNA linear PRI 11-JAN-1997  

GI:1458119  

iens (human)  

iens (human)  

a) Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  

b) Eutheria; Primates; Catarrhini; Homnidae; Homo.  

s 1 to 1876  

K.S.1, Ellis,S.B., Berckhan,K.J., Urrutia,A.,  

Rodriguez,L.E., Johnson,E.C., Vellicelebi,G. and Harpold,M.M.  

The structure of human neuronal alpha 2-alpha 4,beta 2,mu  

nicotinic acetylcholine receptor subunits and functional  

domains of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and  

beta 4 subunits  

Neurosci. 7 (3), 217-228 (1996)  

s 1 to 1876  

K.U.  

Submission  

db (28-JUN-1996) Kathryn J. Ellis et al, SIBIA Neurosciences,  

5 Coast Blvd. So., La Jolla, CA 92037, USA  

Location/Qualifiers  

1..1876  

/organism="Homo sapiens"  

/mol_type="mRNA"  

/db_xref="taxon:9606"  

/cclone="KAlapha7.3"  

/cell_line="IMR32"  

/cell_type="neuroblastoma"  

/cclone_id="SIBIA lambda gt11 library #3 (M. Williams)"  

1..72  

73..1581  

/codon_start=1  

/product="nicotinic acetylcholine receptor alpha7 subunit  

precursor"  

/db_xref="GI:1458120"
```

RESULT 7  
AR055255 1876 bp DNA linear PAT 29-SEP-1999  
LOCUS  
DEFINITION Sequence 7 from patent US 5837489.  
ACCESSION AR055255  
VERSION AR055255.1 GI:5980832  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1876)  
AUTHORS Elliott,K.J., Ellis,S.B. and Harpold,M.M.  
TITLE Human neuronal nicotinic acetylcholine receptor and cells  
transformed with same DNA and mRNA encoding an-subunit of  
JOURNAL Patent: US 5837489-A 7 17-NOV-1998;  
FEATURES  
source Location/Qualifiers  
E COUNT 369 a 553 c 530 g 423 t 1 others  
GIN  
Query Match 27.8%; Score 127; DB 6; Length 1876;  
Best Local Similarity 100.0%; Pred. No. 4.2e-09;  
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 321 GGCCGAGGCGGAGCCCGGCGAGCCGAGAGCTGGAGCGCGCGCTGCTGCAGCT 380  
DB 1 GGCCGAGGCGGAGCCCGGCGAGCCGAGAGCTGGAGCGCGCGCTGCTGCAGCT 60  
QY 381 CCGGAGCTCAACATGCGCTGCTCGCCGAGAGCGCTGCTGCGCTGCGCGCTC 440  
DB 61 CCGGAGCTCAACATGCGCTGCTCGCCGAGAGCGCTGCTGCGCTGCGCGCTC 120  
QY 441 CTGCACG 447  
DB 121 CTGCACG 127

RESULT 8  
AR071403 1876 bp DNA linear PAT 18-FEB-2000  
LOCUS  
DEFINITION Sequence 7 from patent US 5910582.  
ACCESSION AR071403  
VERSION AR071403.1 GI:7222291  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1876)  
AUTHORS Elliott,K.J., Ellis,S.B. and Harpold,M.M.  
TITLE Human neuronal nicotinic acetylcholine receptor compositions and  
methods employing same  
JOURNAL Patent: US 5910582-A 7 08-JUN-1999;  
FEATURES  
source Location/Qualifiers  
BASE COUNT 369 a 553 c 530 g 423 t 1 others  
ORIGIN  
Query Match 27.8%; Score 127; DB 6; Length 1876;  
Best Local Similarity 100.0%; Pred. No. 4.2e-09;  
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 321 GGCCGAGGCGGAGCCCGGCGAGCCGAGAGCTGGAGCGCGCGCTGCTGCAGCT 380  
DB 1 GGCCGAGGCGGAGCCCGGCGAGCCGAGAGCTGGAGCGCGCGCTGCTGCAGCT 60  
QY 381 CCGGAGCTCAACATGCGCTGCTCGCCGAGAGCGCTGCTGCGCTGCGCGCTC 440  
DB 61 CCGGAGCTCAACATGCGCTGCTCGCCGAGAGCGCTGCTGCGCTGCGCGCTC 120  
QY 441 CTGCACG 447  
DB 121 CTGCACG 127

DB 121 CTGCACG 127

RESULT 9  
AR173187 1876 bp DNA linear PAT 17-DEC-2001  
LOCUS  
DEFINITION Sequence 7 from patent US 6303753.  
ACCESSION AR173187  
VERSION AR173187.1 GI:17912678  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1876)  
AUTHORS Elliott,K.J., Ellis,S.B. and Harpold,M.M.  
TITLE Human neuronal nicotinic acetylcholine receptor compositions and  
methods employing same  
JOURNAL Patent: US 6303753-A 7 16-OCT-2001;  
FEATURES  
source Location/Qualifiers  
BASE COUNT 369 a 553 c 530 g 423 t 1 others  
ORIGIN  
Query Match 27.8%; Score 127; DB 6; Length 1876;  
Best Local Similarity 100.0%; Pred. No. 4.2e-09;  
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 321 GGCCGAGGCGGAGCCCGGCGAGCCGAGAGCTGGAGCGCGCGCTGCTGCAGCT 380  
DB 1 GGCCGAGGCGGAGCCCGGCGAGCCGAGAGCTGGAGCGCGCGCTGCTGCAGCT 60  
QY 381 CCGGAGCTCAACATGCGCTGCTCGCCGAGAGCGCTGCTGCGCTGCGCGCTC 440  
DB 61 CCGGAGCTCAACATGCGCTGCTCGCCGAGAGCGCTGCTGCGCTGCGCGCTC 120  
QY 441 CTGCACG 447  
DB 121 CTGCACG 127

RESULT 10  
AR224030 1876 bp mRNA linear PAT 26-SEP-2002  
LOCUS  
DEFINITION Sequence 11 from patent US 6440681.  
ACCESSION AR224030  
VERSION AR224030.1 GI:23332638  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1876)  
AUTHORS Elliott,K.J., Ellis,S.B. and Harpold,M.M.  
TITLE Methods for identifying agonists and antagonists for human neuronal  
nicotinic acetylcholine receptors  
JOURNAL Patent: US 6440681-A 11 27-AUG-2002;  
FEATURES  
source Location/Qualifiers  
BASE COUNT 369 a 553 c 530 g 423 t 1 others  
ORIGIN  
Query Match 27.8%; Score 127; DB 6; Length 1876;  
Best Local Similarity 100.0%; Pred. No. 4.2e-09;  
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 321 GGCCGAGGCGGAGCCCGGCGAGCCGAGAGCTGGAGCGCGCGCTGCTGCAGCT 380  
DB 1 GGCCGAGGCGGAGCCCGGCGAGCCGAGAGCTGGAGCGCGCGCTGCTGCAGCT 60  
QY 381 CCGGAGCTCAACATGCGCTGCTCGCCGAGAGCGCTGCTGCGCTGCGCGCTC 440  
DB 61 CCGGAGCTCAACATGCGCTGCTCGCCGAGAGCGCTGCTGCGCTGCGCGCTC 120



Db 14736 AGCCGAGAGCCCGCTCGGTGAGAGACTGGGGGTGAGAGTCCCGAGCGTAAACCCAGCGCC 14677

QY 121 GGGAGTACTCTCCGCTCACACTCGGGCTGAGTCTCCCTGGTGGCCCGGAGACGCTGG 180

Db 14676 GGGAGTACTCTCCGCTCACACTCGGGCTGAGTCTCCCTGGTGGCCCGGAGACGCTGG 14617

QY 181 CCCGGGCTGAGAGGATGCGGGGCGGGAGCGGGGCGGGGCTCGTCACTGTGAG 240

Db 14616 CCCGGGCTGAGAGGATGCGGGGCGGGAGCGGGGCGGGGCTCGTCACTGTGAG 14557

QY 241 AGGCGCGCGGGGCGGGGCGGGGCGGGGCGGGGCTCGTCACTGTGAG 300

Db 14556 AGGCGCGCGGGGCGGGGCGGGGCGGGGCGGGGCTCGTCACTGTGAG 14497

QY 301 GAGCGGCGAGGTGCTCTGTGGCGGAGCGGCGGGCGGGGCGGGGCTCGTCACTGTGAG 360

Db 14496 GAGCGGCGAGGTGCTCTGTGGCGGAGCGGCGGGCGGGGCGGGGCTCGTCACTGTGAG 14437

QY 361 CGGCGCGGCTCGCTGAGCTCCGGGAGTCAACATGCGCTGCTCGCGGAGGAGGCTGGC 420

Db 14436 CGGCGCGGCTCGCTGAGCTCCGGGAGTCAACATGCGCTGCTCGCGGAGGAGGCTGGC 14377

QY 421 TGGCGCTGCGCGGCTGCTCTGAGCGGTAAAGCCAC 457

Db 14376 TGGCGCTGCGCGGCTGCTCTGAGCGGTAAAGCCAC 14340

RESULT 3  
AF029837 392 bp DNA linear PRI 16-DEC-1998  
LOCUS Homo sapiens alpha 7 neuronal nicotinic receptor gene, promoter  
DEFINITION  
ACCESSION AF029837  
VERSION AF029837.1 GI:3757792  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS Gault, J., Robinson, M., Berger, R., Drebing, C., Logel, J., Hopkins, K., Moore, T., Jacobs, S., Meriwether, J., Choi, M. J., Kim, E. J., Walton, K., Bulting, K., Davis, A., Breese, C., Freedman, R., and Leonard, S.  
TITLE Genomic organization and partial duplication of the human alpha7 neuronal nicotinic acetylcholine receptor gene (CHRNA7)  
JOURNAL MEDLINE 99000837  
PUBMED 9782083  
REFERENCE 2 (bases 1 to 392)  
AUTHORS Leonard, S., Gault, J., Logel, J., Drebing, C., Robinson, M., Berger, R., Breese, C., Davis, A., Hopkins, J., and Freedman, R.  
TITLE Direct Submission  
JOURNAL Submitted (15-OCT-1997) Psychiatry, University of Colorado Health Sciences Center, 4200 E. 9th Ave., Denver, CO 80262, USA  
FEATURES  
source  
1..392  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15q14: between D15S1031 and D15S1360"  
/clone="CHRNA7-P"

Promoter 1..392  
BASE COUNT 51 a 122 c 180 g 39 t  
ORIGIN

Query Match 85.8%; Score 392; DB 9; Length 392;  
Best Local Similarity 100.0%; Pred. No. 2.8e-46;  
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGCGAAGGAGAGAGCTGGCTTGGGCGAGCCCTGGCTGGCGAGAGCGCG 60  
Db 1 AGAAGCGAAGGAGAGAGCTGGCTTGGGCGAGCCCTGGCTGGCGAGAGCGCG 60

QY 61 AGGCCGAGAGCCCGCTCGGTGAGAGACTGGGGGTGAGAGTCCCGAGCGTAAACCCAGCGCC 120

Db 61 AGGCCGAGAGCCCGCTCGGTGAGAGACTGGGGGTGAGAGTCCCGAGCGTAAACCCAGCGCC 120

QY 121 GGGAGTACTCTCCGCTCACACTCGGGCTGAGTCTCCCTGGTGGCCCGGAGACGCTGG 180

Db 121 GGGAGTACTCTCCGCTCACACTCGGGCTGAGTCTCCCTGGTGGCCCGGAGACGCTGG 180

QY 121 GGGAGTACTCTCCGCTCACACTCGGGCTGAGTCTCCCTGGTGGCCCGGAGACGCTGG 180

Db 121 GGGAGTACTCTCCGCTCACACTCGGGCTGAGTCTCCCTGGTGGCCCGGAGACGCTGG 180

QY 181 CCCGGGCTGAGAGGATGCGGGGCGGGAGCGGGGCGGGGCTCGTCACTGTGAG 240

Db 181 CCCGGGCTGAGAGGATGCGGGGCGGGAGCGGGGCGGGGCTCGTCACTGTGAG 240

QY 241 AGGCGCGCGGGGCGGGGCGGGGCGGGGCGGGGCTCGTCACTGTGAG 300

Db 241 AGGCGCGCGGGGCGGGGCGGGGCGGGGCGGGGCTCGTCACTGTGAG 300

QY 301 GAGCGGCGAGGTGCTCTGTGGCGGAGCGGAGCCCGGAGAGCGGAGAGCTGTGAG 360

Db 301 GAGCGGCGAGGTGCTCTGTGGCGGAGCGGAGCCCGGAGAGCGGAGAGCTGTGAG 360

QY 361 CGGCGCGGCTCGCTGAGCTCCGGGAGTCAAC 392

Db 361 CGGCGCGGCTCGCTGAGCTCCGGGAGTCAAC 392

RESULT 4  
AX346920 5304 bp DNA linear PAT 01-FEB-2002  
LOCUS Sequence 1991 from Patent WO0200928.  
DEFINITION AX346920  
ACCESSION AX346920  
VERSION AX346920.1 GI:18494806  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.  
TITLE Diagnosis of diseases associated with the immune system  
JOURNAL Patent: WO 0200928-A 1991 03-JAN-2002;  
Epigenomics AG (DE)  
FEATURES  
source  
1..5304  
/location/Qualifiers  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 1480 a 152 c 1219 g 2453 t  
ORIGIN

Query Match 69.7%; Score 318.6; DB 6; Length 5304;  
Best Local Similarity 81.5%; Pred. No. 2.9e-36;  
Matches 369; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 1 AGAAGCGAAGGAGAGAGTGAAGCTGGCTGGTGGAGAGCCCTGGCTGGCGAGAGCGCG 60  
Db 4609 AGAAGCGAAGGAGAGAGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4668

QY 61 AGGCCGAGAGCCCGCTCGGTGAGAGACTGGGGGTGAGAGTCCCGAGCGTAAACCCAGCGCC 120  
Db 4669 AGGCCGAGAGCCCGCTCGGTGAGAGACTGGGGGTGAGAGTCCCGAGCGTAAACCCAGCGCC 120

QY 121 GGGAGTACTCTCCGCTCACACTCGGGCTGAGTCTCCCTGGTGGCCCGGAGACGCTGG 180  
Db 4729 GGGAGTACTCTCCGCTCACACTCGGGCTGAGTCTCCCTGGTGGCCCGGAGACGCTGG 180

QY 181 CCCGGGCTGAGAGGATGCGGGGCGGGAGCGGGGCGGGGCTCGTCACTGTGAG 240  
Db 4789 TCCGGGCTGAGAGGATGCGGGGCGGGAGCGGGGCGGGGCTCGTCACTGTGAG 240

QY 241 AGGCGCGCGGGGCGGGGCGGGGCGGGGCGGGGCTCGTCACTGTGAG 300  
Db 4849 AGGCGCGCGGGGCGGGGCGGGGCGGGGCGGGGCTCGTCACTGTGAG 300

Anderson, S., Barna, N., Baetien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Glinde, S., Gordon, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Margulis, N., Mathews, L., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L., Mihova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Noth, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seman, S., Severy, P., Spencer, B., Strange-Thomson, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Teefaye, S., Theodore, D., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (10-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 10, 2002 this sequence version replaced gi:18693427.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information  
Center project name: LI0937  
Center clone name: 636\_P\_14

----- Location/Qualifiers  
1. 227506  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15"  
/clone="RP11-636P14"  
/clone\_lib="RPCI-11 Human Male BAC"  
complement(118..329)  
/rpt\_family="MER20"  
repeat\_region  
424..714  
/rpt\_family="AluSg"  
466..476  
/note="<30 qual SNGL region"  
486..508  
/note="<30 qual SNGL region"  
516..520  
/note="<30 qual SNGL region"  
598..602  
/note="<30 qual SNGL region"  
611..617  
/note="<30 qual SNGL region"  
636..640  
/note="<30 qual SNGL region"  
654..658  
/note="<30 qual SNGL region"  
717..1656  
/rpt\_family="L1MEC"  
2145..3397  
/rpt\_family="L1MC1"  
3398..3738  
/rpt\_family="MER1B"  
3743..6334  
/rpt\_family="L1MC1"  
6339..6468  
repeat\_region

repeat\_region /rpt\_family="AluSg/x" 6478..6708  
repeat\_region /rpt\_family="L1MC1" complement(6716..6755)  
repeat\_region /rpt\_family="L1M4" 6776..6942  
repeat\_region /rpt\_family="(TN)n" 6943..9212  
repeat\_region /rpt\_family="L1MC1" 11491..11645  
repeat\_region /rpt\_family="MIR3" complement(11708..11810)  
repeat\_region /rpt\_family="MIR" 14517..14555  
repeat\_region /rpt\_family="GC-rich" 14571..14598  
repeat\_region /rpt\_family="CCCCCG)n" 16447..16557  
repeat\_region /rpt\_family="MIR" 16631..16665  
repeat\_region /rpt\_family="(A)n" 16777..16944  
repeat\_region /rpt\_family="L1MEd" 17266..17585  
repeat\_region /rpt\_family="AluSx" complement(17870..18001)  
repeat\_region /rpt\_family="L1M4" complement(18012..18749)  
repeat\_region /rpt\_family="L1M4" complement(18769..19521)  
repeat\_region /rpt\_family="L1MC/D" complement(19524..19693)  
repeat\_region /rpt\_family="L1" complement(19700..20197)  
repeat\_region /rpt\_family="L1" complement(20198..20642)  
repeat\_region /rpt\_family="MIR1C" complement(20643..20885)  
repeat\_region /rpt\_family="L1" complement(21490..21776)  
repeat\_region /rpt\_family="L1" complement(21844..22045)  
repeat\_region /rpt\_family="L1MDa" complement(22650..22976)  
repeat\_region /rpt\_family="L1MDa" complement(22997..23293)  
repeat\_region /rpt\_family="AluSx" complement(23562..23716)  
repeat\_region /rpt\_family="L1ME3" complement(23862..24071)  
repeat\_region /rpt\_family="L1MC5" complement(24072..24373)  
repeat\_region /rpt\_family="AluSg" complement(24374..24627)  
repeat\_region /rpt\_family="L1MC5" complement(25097..25344)  
repeat\_region /rpt\_family="L1MC5" complement(25670..25805)  
repeat\_region /rpt\_family="HAL1b" complement(25834..26046)  
repeat\_region /rpt\_family="HAL1b" complement(26456..26763)  
Query Match 99.6%; Score 455.4; DB 9; Length 227506;  
Best Local Similarity 99.8%; Pred. No. 4.9e-56;  
Matches 456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGAAGCGAAGGAGAGGTAGAGCTTGAGCTTGAGCAAGCCCTTGAGCCAGAGGCGG 60  
DB 14796 AGAAGCGAAGGAGAGGTAGAGCTTGAGCTTGAGCAAGCCCTTGAGCCAGAGGCGG 14737  
QY 61 AGCCCGAAGAGCCCGCTCGGTGGAGACTGGGGGGGGAGGTGCGCGGAGCGTACCGAGCGCC 120



```

repeat_region /rpt family="MLT1B-int"
                complement(15795..16058)
                /rpt family="L1McA"
repeat_region complement(16601..16727)
                /rpt family="L1McA"
repeat_region 16903..16937
                /rpt family="AT rich"
repeat_region complement(17414..17735)
                /rpt family="AluSx"
repeat_region complement(17795..18092)
                /rpt family="AluO"
repeat_region 18479..18667
                /rpt family="L1M4"
repeat_region 18684..18791
                /rpt family="L1M4"
repeat_region complement(18855..19213)
                /rpt family="L1MEC"
repeat_region complement(19250..19565)
                /rpt family="L1MC1"
repeat_region complement(19640..21583)
                /rpt family="L1MD2"
repeat_region 21593..22179
                /rpt family="L1TR12"
repeat_region complement(22180..23605)
                /rpt family="L1MD2"
repeat_region complement(23627..23725)
                /rpt family="L1MEC"
repeat_region complement(23725..23794)
                /rpt family="L1MEC"
repeat_region complement(24781..24939)
                /rpt family="MIR"
repeat_region 24977..25258
                /rpt family="AluSx"
repeat_region 25622..25908
                /rpt family="AluY"
repeat_region 25916..26033
                /rpt family="MLT1I"
repeat_region 26318..26549
                /rpt family="Charliet1"
repeat_region 26534..26668
                /rpt family="Charliet1"

```

Query Match	99.6%	Score 455.4	DB 9	Length 203493
Best Local Similarity	99.8%	Pred. No. 5.1e-56		
Matches 456	Conservative 0	Mismatch 1	Indels 0	Gaps 0
1	AGAACCCAGAGGAGTAGAGCTTGCTGGAGAGCCCTTGAGAGCTTGAGCTGGCCAGAGGCGC	60		
Db	128648	AGAAACGCAAGGAGAGTAGTAGAGCTTGCTGGAGAGCCCTTGAGCTTGAGCTGGCCAGAGGCGC	128707	
Qy	61	AGGCCAGAGACCCCGCTCCGCTGGAGAGATTGGGGGTGAGGTGCCCGAGCGTATCCAGCGCC	120	
Db	128708	AGGCCAGAGACCCCGCTCCGCTGGAGAGATTGGGGGTGAGGTGCCCGAGCGTATCCAGCGCC	128767	
Qy	121	GGGAGTACCTCCCGCTCACACCTCGGGGTGAGATTCCCTGGGGTGGCCCGAGACGCTGG	180	
Db	128768	GGGAGTACCTCCCGCTCACACCTCGGGGTGAGATTCCCTGGGGTGGCCCGAGACGCTGG	128827	
Qy	181	CCGGGCTTGGAGGAGATGGCGGGGCGGGAGCGGGGCGGGGCTCGTCACTGGAG	240	
Db	128828	CCGGGCTTGGAGGAGATGGGGGGGCGGGGCGGGGCGGGGCTCGTCACTGGAG	128887	
Qy	241	AGCGCGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGCTCTTTAAAGGCGCGAGCC	300	
Db	128888	AGCGCGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGCTCTTTAAAGGCGCGAGCC	128947	
Qy	301	GAGCGCGAGGTCTCTTGTGCGCCGAGGCGGAGGCCCGGGCGACAGCCGAGAGCTGGAG	360	
Db	128948	GAGCGCGAGGTCTCTTGTGCGCCGAGGCGGAGGCCCGGGCGACAGCCGAGAGCTGGAG	129007	
Qy	361	CGGCGCGGCTCGCTGAGCTCCGGGAGCTCAATGCGCTGCTCCCGGAGAGCGTCTGGC	420	
Db	129008	CGGCGCGGCTCGCTGAGCTCCGGGAGCTCAATGCGCTGCTCCCGGAGAGCGTCTGGC	129067	

QY	421	TGGCGCTGACCGCGTGCCTGACGAGGAAAGCCAC	457
Db	129068	TGGCGCTGACCGCGTGCCTGACGAGGAAAGCCAC	129104
RESULT 2	AC079969/c	227506 bp	DNA linear
LOCUS	AC079969/c	227506 bp	DNA linear
DEFINITION	Homo sapiens chromosome 15, clone RP11-636P14, complete sequence.		
ACCESSION	AC079969		
VERSION	AC079969.5	GI:20522205	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 227506)		
TITLE	Birren, B., Linton, L., Nusbaum, C. and Lander, E.		
JOURNAL	Homo sapiens chromosome 15, clone RP11-636P14		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 227506)		
TITLE	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Barna, N., Bastien, V., Beda, F., Boguslavsky, L., Boukhalter, B., Brown, A., Burbett, G., Campopiano, A., Castle, A., Chepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J. S., Dodge, S., Ferris, P., Fitzhugh, W., Gage, D., Galagan, J., Galdyna, S., Ginde, S., Goyette, M., Graham, I., Grand-Pierre, N., Hagos, B., Hatford, A., Horton, L., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Larocque, K., Lamazares, R., Landers, T., Lechoczy, J., Levine, R., Liu, C., Liu, G., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K., McSheeters, R., Meldrum, J., Menue, L., Mihova, T., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Piere, N., Pisan, C., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Sougnaz, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirelli, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	3 (bases 1 to 227506)		
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Chepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Galdyna, S., Ginde, S., Goyette, M., Graham, I., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R., Landers, T., Lechoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menue, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicoll, R., Nobbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Piere, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tophan, K., Travers, M., Trivis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	4 (bases 1 to 227506)		
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,		



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2003, 05:26:09 ; Search time 2500 Seconds

(without alignments)  
7478.280 Million cell updates/sec

Title: US-08-956-518A-94

Perfect score: 457

Sequence: 1 AGAAGCAGAGGAGAGGTAG.....CTCCTGACCGTAAAGCCAC 457

Scoring table:

IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

2888711 segs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: GenEmbl:  
1: gb\_ba:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hg\_hum:\*  
31: em\_hg\_inv:\*  
32: em\_hg\_other:\*  
33: em\_hg\_mus:\*  
34: em\_hg\_pln:\*  
35: em\_hg\_rtd:\*  
36: em\_hg\_mam:\*  
37: em\_hg\_vtc:\*  
38: em\_sy:\*  
39: em\_hgo\_hum:\*  
40: em\_hgo\_mus:\*  
41: em\_hgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	455.4	99.6	203493	9 AC058803	AC058803 Homo sapi
2	455.4	99.6	227506	9 AC079969	AC079969 Homo sapi
3	392	85.8	332	9 AF029837	AF029837 Homo sapi
4	318.6	49.7	5304	6 AX346920	AX346920 Sequence
5	221.8	48.5	5304	6 AX346921	AX346921 Sequence
6	133.8	29.3	2107	9 AF486623	AF486623 Macaca mu
7	127	27.8	1876	6 AR055255	AR055255 Sequence
8	127	27.8	1876	6 AR071403	AR071403 Sequence
9	127	27.8	1876	6 AR173187	AR173187 Sequence
10	127	27.8	1876	6 AR224030	AR224030 Sequence
11	127	27.8	1876	6 AR282833	AR282833 Sequence
12	127	27.8	1876	6 AX719088	AX719088 Sequence
13	127	27.8	1876	9 HS062436	HS062436 Human n1cot
14	121.6	26.6	2087	9 HSAARA7A	X70297 H. sapiens m
15	94	20.6	1810	9 BC037571	BC037571 Homo sapi
16	82.2	18.0	219952	2 AC084804	AC084804 Mus muscu
17	81.6	17.9	1393	11 PM1H12G	PM1H12G Penicilli
18	81.2	17.8	141534	2 AP003685	AP003685 Oryza sat
19	80.2	17.5	155680	2 AC120887	AC120887 Oryza sat
20	78.8	17.2	81767	2 AC021929	AC021929 Homo sapi
21	78.6	17.2	64207	2 AC024002	AC024002 Homo sapi
22	78.6	17.2	146743	2 AP003518	AP003518 Oryza sat
23	78.2	17.1	9711	11 PM1285B	PM1285B Penicilli
24	78	17.0	78220	2 AC024266	AC024266 Homo sapi
25	77.8	17.1	65523	2 AC023212	AC023212 Homo sapi
26	77.4	16.9	885	11 PM7F8G	PM7F8G Penicilli
27	76.8	16.8	63087	2 AC023446	AC023446 Homo sapi
28	76.6	16.8	113141	2 AC093919	AC093919 Oryza sat
29	76	16.6	129506	2 AC136100	AC136100 Rattus no
30	75.4	16.5	120091	2 AC012684	AC012684 Homo sapi
31	75.4	16.5	144979	2 AC016280	AC016280 Homo sapi
32	74.8	16.4	87283	2 AC022856	AC022856 Homo sapi
33	74.6	16.3	149454	2 AC013648	AC013648 Homo sapi
34	74.4	16.3	976	11 PM12D6G	PM12D6G Penicilli
35	74	16.2	1052	11 PM2H12B	PM2H12B Penicilli
36	73.8	16.1	11305	8 OSAS55061	OSAS55061 Oryza sat
37	73.8	16.1	159790	2 AP004786	AP004786 Oryza sat
38	73.8	16.1	220389	2 BX324177	BX324177 Danio rer
39	73.6	16.1	75628	2 AC021793	AC021793 Homo sapi
40	73.6	16.1	164520	2 AC020738	AC020738 Homo sapi
41	73.4	16.1	303091	2 AC084799	AC084799 Mus muscu
42	73	16.0	1094	11 PM7G11B	PM7G11B Penicilli
43	73	16.0	167624	2 AC143286	AC143286 Macaca mu
44	72.8	15.9	76734	2 AC023218	AC023218 Homo sapi
45	72.4	15.8	114639	9 AC100810	AC100810 Homo sapi

## ALIGNMENTS

RESULT 1  
AC058803  
LOCUS AC058803 203493 bp DNA linear PRI 11-DEC-2001  
DEFINITION Homo sapiens chromosome 15, clone RP11-717124, complete sequence.  
AC058803  
ACCESSION AC058803  
VERSION AC058803.5 GI:17488731  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 203493)  
AUTHORS Birren B., Linton L., Nusbaum C. and Lander E.  
TITLE Homo sapiens chromosome 15, clone RP11-717124  
JOURNAL Unpublished

QY 381 CCGGACTCAACATGCGCTGCGCCGGAGGCGTGGCGCTGGCCGCGCTC 440  
|||||  
Db 61 CCGGACTCAACATGCGCTGCTCGCGGAGGCGCTGGCTGGCGCTGGCGCTC 120  
|||||  
QY 441 CTGCACG 447  
|||||  
Db 121 CTGCACG 127

RESULT 6  
US-08-660-451A-11  
Sequence 11, Application US/08660451A  
Parent No. 6524789

GENERAL INFORMATION:  
APPLICANT: Eliott, Kathryn J.  
APPLICANT: Harpold, Michael M.  
TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE  
TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: CA  
COUNTRY: USA  
ZIP: 92101-2926

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/660,451A  
FILING DATE: June 7, 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/484,722  
FILING DATE: 06/07/95

ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-9370B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-238-0999  
TELEFAX: 619-238-0062  
TELEX:

INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1876 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 73...1581  
OTHER INFORMATION: alpha7 human neuronal nicotinic  
OTHER INFORMATION: acetylcholine receptor  
NAME/KEY: 5'UTR  
LOCATION: 1...72  
OTHER INFORMATION:  
NAME/KEY: 3'UTR  
LOCATION: 1582...1876  
OTHER INFORMATION:

US-08-660-451A-11

Query Match 27.8%; Score 127; DB 4; Length 1876;  
Best Local Similarity 100.0%; Pred. No. 1,1e-15;

Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 321 GCGCGCAGGCGCGAGCGCCGGGCGACAGCCAGACGTGGAGCCGCGCTCGCTGCAGCT 380  
|||||  
Db 1 GCGCGCAGGCGCGAGCGCCGGGCGACAGCCAGACGTGGAGCGCGCCGCGCTCGCTGCAGCT 60  
|||||  
QY 381 CCGGACTCAACATGCGCTGCGCCGGAGGCGTGGCTGGCGCTGGCCGCGCTC 440  
|||||  
Db 61 CCGGACTCAACATGCGCTGCTCGCGGAGGCGCTGGCTGGCGCTGGCGCTC 120  
|||||

QY 441 CTGCACG 447  
|||||  
Db 121 CTGCACG 127

RESULT 7  
US-08-318-837-1/c  
Sequence 1, Application US/0831837  
Patent No. 5981277

GENERAL INFORMATION:  
APPLICANT: FRANSSEN, LUCIA; DEVOS, KATHLEEN; VAN DE VOORDE,  
APPLICANT: ANDRE; VAN HEVERSWYN, HUGO  
TITLE OF INVENTION: NEW POLYPEPTIDES AND PEPTIDES, NUCLEIC ACID  
TITLE OF INVENTION: CODING FOR THEM, AND THEIR USE IN THE FIELD OF TUMOR THERAPY  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIERMAN AND MUSERLIAN  
STREET: 600 THIRD AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10016

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/318,837  
FILING DATE: 13-OCT-1994  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP 93/01022  
FILING DATE: 28-APR-1993

ATTORNEY/AGENT INFORMATION:  
NAME: CHARLES A. MUSERLIAN  
REGISTRATION NUMBER: 19,663  
REFERENCE/DOCKET NUMBER: 410,007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 661-8000  
TELEFAX: (212) 661-8002

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2150 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Human  
TISSUE TYPE: Spleen tissue from healthy human  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1810...1982  
OTHER INFORMATION:  
NAME/KEY: misc\_feature  
LOCATION: 543

Db 121 CTGCACG 127

; sequence 7, Application US/09217345  
; Patent No. 6303753

GENERAL INFORMATION:  
APPLICANT: E11104

APPLICANT: Ellis, Steven B.

APPLICANT: Harpold, Michael M.  
TITLE OF INVENTION: METHODS FOR IDENTIFYING AGONISTS AND

TITLE OF INVENTION: ANTAGONISTS FOR HUMAN NEURONAL  
NICOTINIC ACETYL CHOLINE RECEPTORS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brown, Martin, Halper & McClain

STREET: 1660 UN  
CITY: San Diego

STATE: CA

COUNTRY: USA  
ZIP: 92101

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5 1/4 inch disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: 70/60,402-200

AFFILIATION NUMBER: US/08/487,596  
 FILING DATE: 07-JUN-1995

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO US94/02447  
FILING DATE: 08 MAR 1994

FILING DATE: 08-MAR-1994  
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/149,503  
FILING DATE: 08-NOV-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/028,031  
FILING DATE: 08-MAR-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IIS 07/938 154

FILING DATE: 30-NOV-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/504,455

FILING DATE: 03-APR-1990  
ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie L.

REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-995

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-238-0999

TELEFAX: 619-238-0062

INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1876 base pairs  
TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both  
MOLECULE TYPE: cDNA

FEATURE:  
NAME/KEY: CDS

LOCATION: 73.1581

OTHER INFORMATION: /product= "ALPHA  
08-487-596-11

27 04. 60000 100

```

best Local Similarity 100.0%; Pred. N

```

atches 127; Conservative 0; Mismatches 0

321 GCGCGAGCGCAGGCCGGGCGACAGC

1 GCGCGAGCGCAGGCCCGGCGACAGC

Query Match	27.8%;	Score 127;	DB 2;	Length 1876;
Best Local Similarity	100.0%;	Pred. No. 1.1e-15;		
Matches 127;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

[illegible]

## RESULT 2

US-08-700-636-7  
Sequence 7, Application US/08700636  
Patent No. 5910582

```

GENERAL INFORMATION:
APPLICANT: Elliott, Kathryn J.
APPLICANT: Ellis, Steven B.
APPLICANT: Harpold, Michael M.
TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,636
FILING DATE: 16-JUL-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US
FILING DATE: 08-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Keltner, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9368
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1876 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 73..1581
/JS-08-700-636-7

```

Query March	27.8%	Score 127	DB 2	Length 1876
Best Local Similarity	100.0%	Pred. No. 1.1e-15		
Matches 127	Conservative 0	Mismatches 0	Indels 0	Gaps 0

  

QY	321	GCGCCGACAGCCAGGCCCGGGGAGACCGCGAGCGCGCGCTGCTGCAGCT	380
Db	1	GCGCCGACAGCCAGGCCCGGGGAGACCGTGGAGCGCGCGCGCTGCTGCAGCT	60

Oy	381	CGGGAGCTAA	CTGGCGTCTCGGGGGAGAGCGCTT	AGCTGGCGCGTGGCCGCTCGCTC	440
Db	61	CGGGAGCTAA	CACTGGCTCTCGCGGGAGGCGT	CTGGCTGGCGCGCTCGCTC	120
Oy	441	CTGGACG	447		
Db	121	CTGGACG	127		

RESULT 3

US-08-467-534-1  
; Sequence 7, Application US/08467574  
; Patent No. 6022704

1 GENERAL INFORMATION:  
1 APPLICANT: Elliot, Kathryn J.  
1 APPLICANT: Ellis, Steven B.  
1 APPLICANT: Harpold, Michael M.  
1 TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE  
1 TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME  
1 NUMBER OF SEQUENCES: 12  
1 CORRESPONDENCE ADDRESS:  
1 ADDRESSEE: Brown, Martin, Haller & McClain  
1 STREET: 1660 Union Street  
1 CITY: San Diego  
1 STATE: CA  
1 COUNTRY: USA  
1 ZIP: 92101-2926  
1 COMPUTER READABLE FORM:  
1 MEDIUM TYPE: Diskette  
1 COMPUTER: IBM Compatible  
1 OPERATING SYSTEM: DOS  
1 SOFTWARE: FASTSEQ Version 1.5  
1 CURRENT APPLICATION DATA:  
1 APPLICATION NUMBER: US/08/467,574  
1 FILING DATE: June 5, 1995  
1 CLASSIFICATION: 536  
1 PRIOR APPLICATION DATA:  
1 APPLICATION NUMBER: US 08/028,031  
1 FILING DATE: March 8, 1993  
1 ATTORNEY/AGENT INFORMATION:  
1 NAME: Seidman, Stephanie L.  
1 REGISTRATION NUMBER: 33,779  
1 REFERENCE/DOCKET NUMBER: 6362-9949  
1 TELECOMMUNICATION INFORMATION:  
1 TELEPHONE: 619-238-0999  
1 TELEFAX: 619-238-0062  
1 TELEX:  
1 INFORMATION FOR SRO ID NO: 7:  
1 SEQUENCE CHARACTERISTICS:  
1 LENGTH: 1876 base pairs  
1 TYPE: nucleic acid  
1 STRANDEDNESS: both  
1 TOPOLOGY: both  
1 MOLECULE TYPE: cDNA  
1 FEATURE:  
1 NAME/KEY: CDS  
1 LOCATION: 73..1581  
1 US-08-467-574-7

	Query Match	27.8%	Score 127	DB 3	Length 1876
	Best Local Similarity	100.0%	Pred. NO. 1.1e-15		
	Matches 127	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	321	GGCCGACAGCGCCGCGGCGACAGCCGACGCTGAGAGCGCGCGCTGCTGCACCT			380
Db	1	GGCGGAGCGCCGACCCCGGCGACAGCCGACGCTGAGAGCGCGCGCTGCTGCACCT			60
QY	381	CCGGGACTCAACATGCGCTGCTCGCCGCGAGAGCGCTTGCGCGCTGAGCGCGCTG			440
Db	61	CCGGGACTCAACATGCGCTGCTCGCCGCGAGAGCGCTTGCGCGCGCTGAGCGCGCTG			120
QY	441	CTGCACG 447			

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2003, 08:56:10 ; Search time 68 Seconds  
(without alignments)  
2966.354 Million cell updates/sec

Title: US-08-956-518A-94  
Perfect score: 457  
Sequence: 1 AGAAGCGAAGAGAGAGTAG.....CTCCTGCGACGTAAGCCAC 457

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

1 number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PTCUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	* Match Length	ID	Description
1	127	27.8	1876 2	US-08-466-589-7 Sequence 7, Appl
2	127	27.8	1876 2	US-08-700-636-7 Sequence 7, Appl
3	127	27.8	1876 3	US-08-467-574-7 Sequence 7, Appl
4	127	27.8	1876 4	US-09-217-345-7 Sequence 7, Appl
5	127	27.8	1876 4	US-08-487-596-11 Sequence 11, Appl
6	127	27.8	1876 4	US-08-660-451A-11 Sequence 11, Appl
7	69.2	15.1	2150 2	US-08-318-837-1 Sequence 16, Appl
8	64	14.0	152331 3	US-09-128-155-16 Sequence 16, Appl
9	60.6	13.3	320 3	US-09-165-264-13 Sequence 13, Appl
10	60.2	13.2	319 3	US-09-165-264-8 Sequence 8, Appl
11	58.8	12.9	320 3	US-09-165-264-7 Sequence 7, Appl
12	58.8	12.9	320 3	US-09-165-264-14 Sequence 14, Appl
13	56.8	12.4	320 3	US-09-165-264-11 Sequence 11, Appl
14	56.4	12.3	318 3	US-09-165-264-12 Sequence 12, Appl
15	56.4	12.3	12001 1	US-08-458-568A-11 Sequence 11, Appl
16	56	12.3	1120 3	US-09-030-613-1 Sequence 1, Appl
17	56	12.3	1120 3	US-09-451-905-1 Sequence 1, Appl
18	54.8	12.0	4403765 3	US-09-103-840A-2 Sequence 2, Appl
19	54.2	11.9	22187 4	US-09-499-522-1 Sequence 1, Appl
20	54	11.8	2000 4	US-09-338-907-180 Sequence 180, App
21	54	11.8	2000 4	US-09-218-207-180 Sequence 180, App
22	54	11.8	56516 2	US-08-996-306-1 Sequence 1, Appl
23	54	11.8	56516 3	US-09-338-907-1 Sequence 1, Appl
24	54	11.8	56516 4	US-09-218-207-1 Sequence 1, Appl
25	54	11.8	56520 3	US-09-338-907-179 Sequence 179, App
26	54	11.8	56520 4	US-09-218-207-179 Sequence 179, App
27	53.2	11.6	4411529 3	US-09-103-840A-1 Sequence 1, Appl

C 28	53	11.6	4257 2	US-08-690-473-1 Sequence 1, Appl
C 29	53	11.6	4257 3	US-09-259-821A-1 Sequence 1, Appl
C 30	53	11.6	4257 3	US-08-843-659-1 Sequence 1, Appl
C 31	53	11.6	12001 1	US-08-458-568A-11 Sequence 11, Appl
C 32	52.4	11.5	5228 4	US-09-428-711A-15 Sequence 15, Appl
C 33	52.4	11.5	5228 4	US-09-428-711A-15 Sequence 15, Appl
C 34	52.4	11.5	7218 1	US-08-232-463-14 Sequence 14, Appl
C 35	52.4	11.5	4403765 3	US-09-103-840A-2 Sequence 2, Appl
C 36	51.8	11.3	4403 2	US-08-284-941-1 Sequence 1, Appl
C 37	51.8	11.3	4403 2	US-08-447-642-1 Sequence 1, Appl
C 38	51.8	11.3	4403 3	US-09-236-503-1 Sequence 1, Appl
C 39	51.8	11.3	4403 3	PCT-US93-02147A-1 Sequence 1, Appl
C 40	51.8	11.3	4411529 3	US-09-103-840A-1 Sequence 1, Appl
C 41	51.4	11.2	4060 1	US-08-308-949A-1 Sequence 1, Appl
C 42	50.8	11.1	407 4	US-09-056-556-173 Sequence 173, App
C 43	50.8	11.1	407 4	US-09-072-596-168 Sequence 168, App
C 44	50.4	11.0	1292 3	US-08-483-533-37 Sequence 37, Appl
C 45	50.4	11.0	1292 4	US-09-283-471A-37 Sequence 37, Appl

## ALIGNMENTS

RESULT 1  
US-08-466-589-7  
; Sequence 7, Application US/08466589  
; Patent No. 5837489  
; GENERAL INFORMATION:  
; APPLICANT: Eliot, Kathryn J.  
; APPLICANT: Ellis, Steven B.  
; APPLICANT: Harpold, Michael M.  
; TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE  
; TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brown, Martin, Haller & McClain  
; STREET: 1660 Union Street  
; CITY: San Diego  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92101-2926  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/466,589  
; FILING DATE: June 5, 1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/028,031  
; FILING DATE: March 8, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Seidman, Stephanie L.  
; REGISTRATION NUMBER: 33,779  
; REFERENCE/DOCKET NUMBER: 6362-9950  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-238-0999  
; TELEFAX: 619-238-0062  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1876 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 73..1581  
; US-08-466-589-7